ARTICLE

doi:10.1038/nature15699

Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65

Xi-Ping Huang[1,2]*, Joel Karpiak[3]*, Wesley K. Kroeze[1]*, Hu Zhu[1]†, Xin Chen[4,5]†, Sheryl S. Moy[6], Kara A. Saddoris[6]†, Viktoriya D. Nikolova[6], Martilias S. Farrell[1]†, Sheng Wang[1], Thomas J. Mangano[1,2], Deepak A. Deshpande[7], Alice Jiang[1,2]†, Raymond B. Penn[7], Jian Jin[4,5]†, Beverly H. Koller[8], Terry Kenakin[1], Brian K. Shoichet[3] & Bryan L. Roth[1,2,5]

At least 120 non-olfactory G-protein-coupled receptors in the human genome are 'orphans' for which endogenous ligands are unknown, and many have no selective ligands, hindering the determination of their biological functions and clinical relevance. Among these is GPR68, a proton receptor that lacks small molecule modulators for probing its biology. Using yeast-based screens against GPR68, here we identify the benzodiazepine drug lorazepam as a non-selective GPR68 positive allosteric modulator. More than 3,000 GPR68 homology models were refined to recognize lorazepam in a putative allosteric site. Docking 3.1 million molecules predicted new GPR68 modulators, many of which were confirmed in functional assays. One potent GPR68 modulator, ogerin, suppressed recall in fear conditioning in wild-type but not in GPR68-knockout mice. The same approach led to the discovery of allosteric agonists and negative allosteric modulators for GPR65. Combining physical and structure-based screening may be broadly useful for ligand discovery for understudied and orphan GPCRs.

G-protein-coupled receptors (GPCRs)—the largest family of proteins encoded in the human genome—transduce signals for the most diverse endogenous ligands of any receptor family. Correspondingly, GPCRs are the most productive drug targets, with over 26% of US Food and Drug Administration (FDA)-approved drugs acting primarily through them. Astonishingly, of the 356 non-olfactory GPCRs, about 38% are understudied or 'orphan' receptors whose physiological roles, and often endogenous ligands, remain unknown[1]. Given the central role of GPCRs in physiology and disease, and the high conservation of orphan GPCRs among organisms from worms to humans, understudied and orphan GPCRs are probably functionally and therapeutically important. Indeed, for the few GPCRs deorphanized since 2003 (refs 1, 2 and http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=16), most have newly approved and investigational drugs[1,3]. As with kinases[4], epigenetic proteins[5] and proteases[6], ligands specific for orphan GPCRs will illuminate their biology and provide new areas for therapeutic intervention.

A key impediment to GPCR deorphanization is uncertainty about the proteins through which they signal, making functional assays problematic[1]. This difficulty is increased by the diverse ligands that GPCRs recognize, which range from protons and photons, small neurotransmitters and lipids, to peptides and folded proteins. Thus, generic functional screens are difficult for orphan GPCRs—one neither knows what class of compounds to screen, nor how to screen for it, much less how to demonstrate relevance—thereby explaining the slow progress in determining their roles in signalling and physiology[3].

GPR68 (also known as OGR1) exemplifies both the important roles these understudied and orphan receptors are thought to serve, and our difficulties in illuminating them. Together with GPR4, GPR65 and GPR132, GPR68 belongs to a family of proton-sensing GPCRs[7]. GPR68 couples to several signalling pathways through G_q, G_s, $G_{12/13}$ or $G_{i/o}$ proteins[7–10]. GPR68 is expressed in many tissues and has been implicated in many processes[11–16], but it is most abundant in mouse cerebellum[17] and hippocampus[11] (http://www.brain-map.org/), suggesting yet to be identified roles in brain function. In acidic microenvironments, GPR68 seems to regulate inflammatory processes in airway smooth muscle and other cells[18–20]. Surprisingly, studies with GPR68-knockout mice uncovered only modest changes in these functions[16,21,22]. Although GPR68 has been reported to be activated by a family of isoxazoles[15], their weak activity seems to be nonspecific[23,24] and could not be reproduced (see later). Thus, although GPR68 may have many roles, few of them are well-characterized by knockout and none is known in the central nervous system (CNS), where it is most highly expressed. Like other targets lacking small molecule reagents, GPR68 remains 'pharmacological dark matter'[1].

Here we describe an integrated experimental and computational approach to discover ligands that modulate GPR68. A lead compound that functions as a positive allosteric modulator (PAM) is demonstrated *in vitro* and *in vivo*, providing insights into GPR68 physiology. Application of the same approach found allosteric agonists and negative allosteric modulators for a second understudied GPCR, GPR65, suggesting that the approach may be broadly useful.

[1]Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599-7365, USA. [2]National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7365, USA. [3]Department of Pharmaceutical Chemistry, University of California at San Francisco, Byers Hall, 1700 4th Street, San Francisco, California 94158-2550, USA. [4]Center for Integrative Chemical Biology and Drug Discovery (CICBDD), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7363, USA. [5]Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7360, USA. [6]Department of Psychiatry and Carolina Institute for Developmental Disabilities (CIDD), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7146, USA. [7]Center for Translational Medicine and Department of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA. [8]Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7264, USA. †Present addresses: National Center for Advancing Translational Sciences (NCATs), 9800 Medical Center Drive, Rockville, Maryland 20850, USA (H.Z.); Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7365, USA (X.C.); Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado 80309, USA (K.A.S.); Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA (M.S.F.); Touro College of Osteopathic Medicine, 60 Prospect Avenue, Middletown, New York 10940, USA (A.J.); Department of Structural and Chemical Biology, Department of Oncological Sciences, Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA (J.J.). *These authors contributed equally to this work.

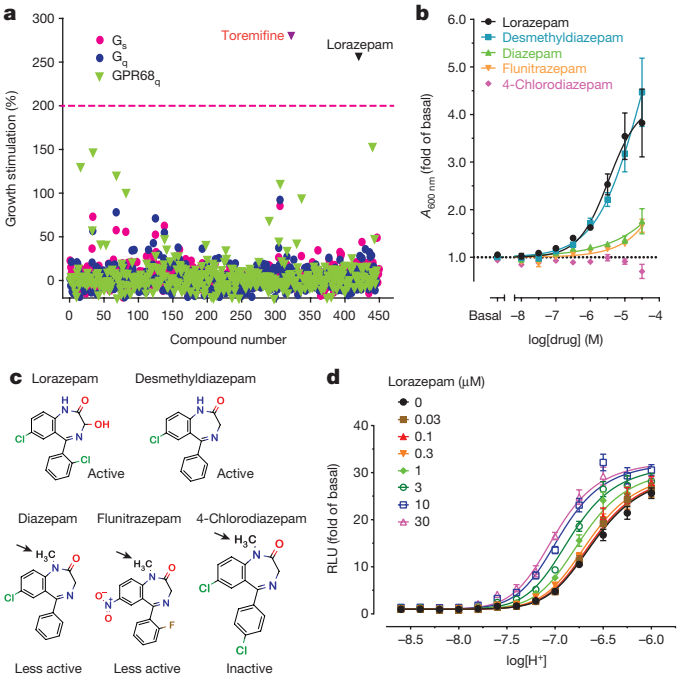


Figure 1 | Lorazepam is a GPR68-positive allosteric modulator. **a**, A library of approved drugs (10 μM) screened with yeast expressing chimaeric G_s, G_q or GPR68 and chimaeric G_q (GPR68$_q$) revealed lorazepam as a true and toremifine as a false positive. **b**, Concentration-dependent stimulation of GPR68 G_q-yeast growth by lorazepam and analogues. **c**, Structures of representative benzodiazepines (arrows denote methyl substituents that reduce GPR68 activity). **d**, Lorazepam is a GPR68-positive allosteric modulator for the agonist proton in the GPR68-mediated cAMP production. RLU, relative luminescence units. Data are mean ± s.e.m. of normalized results (**a**, **b**, **d**, $n = 3$) and concentration–response curves (**b**, **d**) were fit via a four-parameter logistic function (see Methods).

Yeast-based screen reveals GPR68 active compounds

In an initial campaign with 24 selected orphan and understudied GPCRs, we modified a yeast assay system[25] and screened a small library of approved drugs (http://www.nihclinicalcollection.com/ and Supplementary Fig. 1). We confirmed the known activity of short-chain carboxylic acids on the GPR41 and the GPR43 free fatty acid receptors (Extended Data Fig. 1a–d), and that of zinc (Extended Data Fig. 1e) and several other metals (Extended Data Fig. 1f–k) at GPR39. The most notable result was the finding that the benzodiazepine anxiolytic lorazepam was an agonist at GPR68 (Fig. 1).

Lorazepam activated GPR68 signalling, stimulating yeast growth by more than twofold (Fig. 1a). *N*-unsubstituted benzodiazepines were more efficacious than *N*-substituted benzodiazepines (Fig. 1b, c and Supplementary Table 1) and activated the receptor at both pH 6.5 and 7.4 (Extended Data Fig. 1l), with lorazepam most potently shifting the H^+ concentration–response profile (Fig. 1d and Extended Data Fig. 1m–p). The pH-dependence of lorazepam activity suggested that it functions as a PAM of GPR68; lorazepam did not affect the activity of the related receptors GPR4 or GPR65 (Extended Data Fig. 2a, b). When profiled against a panel of CNS targets, lorazepam had substantial activity only at the $GABA_A$ (γ-aminobutyric acid type A) receptor, its therapeutic target (Extended Data Fig. 3).

Modelling the GPR68–lorazepam complex

Little improvement in activity or selectivity was achieved by testing lorazepam analogues. This observation, and the potent $GABA_A$ receptor activity of the drug, led us to seek specific, optimizable molecules from computational docking screens of multi-million molecule libraries (Fig. 2).

We generated 407 homology 3D models for GPR68 templated on the CXCR4 structure (29% sequence identity, Extended Data Fig. 2f), and these were expanded by another 2,900 models using elastic network modelling, which sampled backbone and loop conformations. Against each of the 3,307 models, we computationally docked the active benzodiazepines, more than 440 inactive compounds from the National Clinical Collection (NCC; http://nihsmr.evotec.com/evotec/sets) library, and 176 property-matched decoy molecules[26]. In each model, five candidate allosteric sites were docked against (Extended Data Fig. 2g), based on the binding regions of aminergic GPCRs, the peptide and antagonist sites of CXCR4, and the muscarinic receptor allosteric site. Iterative cycles of modelling and optimization (Fig. 2b–e) attempted to capture two aspects of ligand binding. First, the activity of

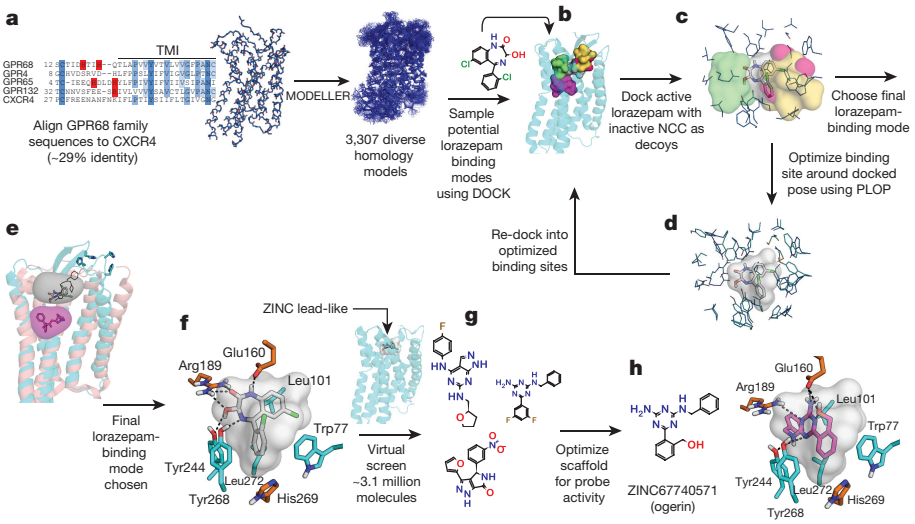


Figure 2 | Virtual screening workflow and predicted location of GPR68 allosteric site. **a**, Sequence alignment of GPR68, GPR4, GPR65 and GPR132 to CXCR4 (details in Extended Data Fig. 2e). **b**, Docking of lorazepam and NCC library to five distinct binding sites (details in Extended Data Fig. 2f). **c**, Models evaluated by their favourable ranking of lorazepam versus decoy molecules. **d**, Optimizing the most favourable lorazepam binding mode. **e**, Optimized lorazepam orientation (grey stick) in GPR68 (cyan ribbon) and M_2 muscarinic receptor (salmon ribbon; Protein Data Bank (PDB) code 4MQT) with allosteric site (grey) and orthosteric site (quinuclidinyl benzilate, magenta). **f**, Lorazepam in its predicted orientation and interactions. **g**, Virtual screen of ZINC subset (~3.1 million molecules) to identify predicted hits. **h**, ZINC67740571 (magenta stick) in its predicted orientation and interactions in GPR68.

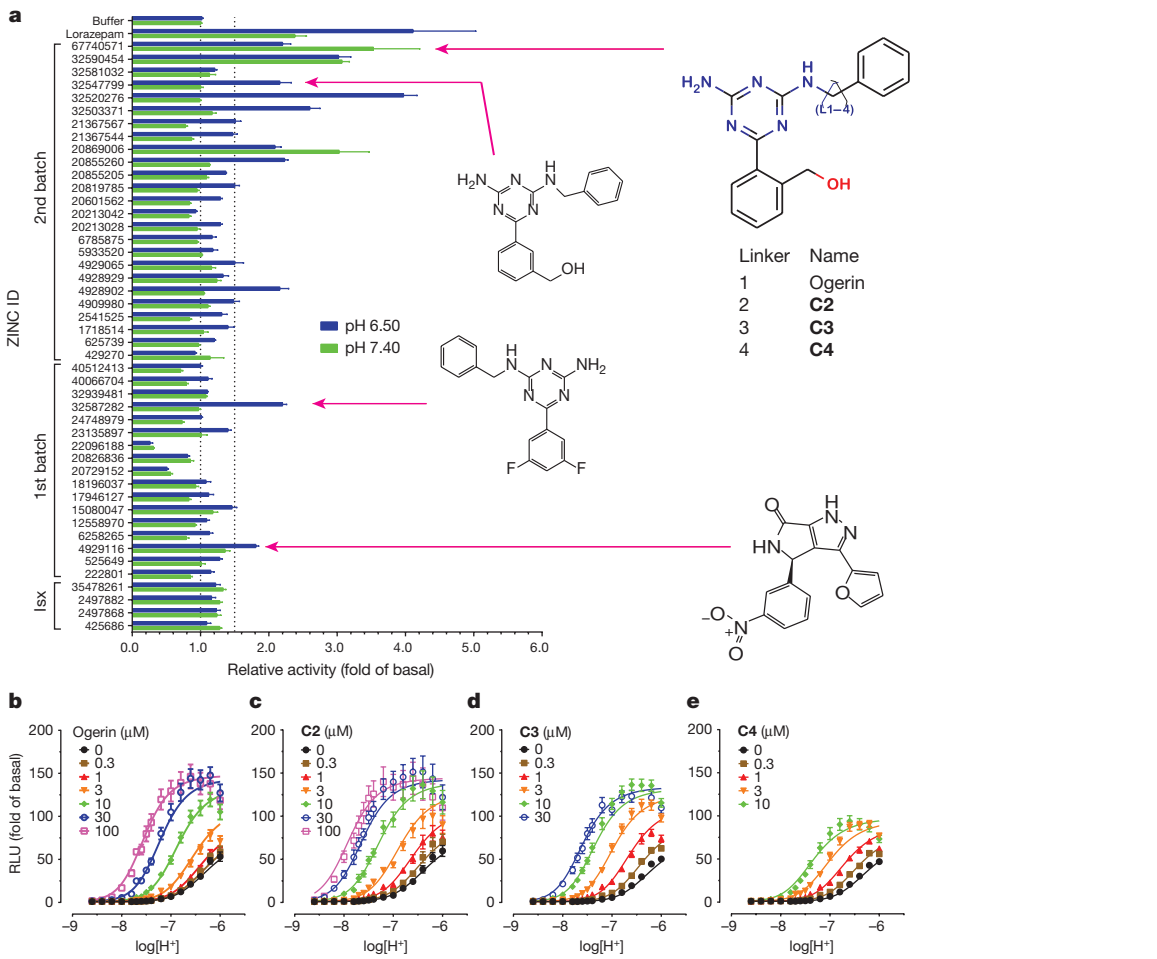


Figure 3 | Identification, characterization and optimization of GPR68-positive allosteric modulators. a, Normalized results of GPR68-mediated cAMP production for selected compounds (ZINC database numbers) are shown; data represent mean ± s.e.m. ($n = 4$–34 measurements) at 10 μM for pH 7.40 and 6.50. Compounds were grouped into a first batch from the first round of virtual docking, and a second batch from the second round of docking. Compounds labelled Isx are isoxazole analogues. Lead compounds ZINC32587282, ZINC4929116, ZINC67740571 (ogerin), its isomer (ZINC32547799) and analogues (**C2**, **C3** and **C4**) with different lengths of linkers, are highlighted. **b**–**e**, Concentration–response curves of normalized data (mean ± s.e.m.; $n = 4$) for ogerin (**b**), **C2** (**c**), **C3** (**d**) and **C4** (**e**) are shown to illustrate the allosteric potentiation of proton and analysed using a standard operational allosteric model. Allosteric parameters are summarized in Supplementary Table 8, and curve-fitting details are in Methods.

the benzodiazepines as PAMs, and second, the role of histidine residues 17, 84, 169 and 269, which are thought to interact with one another in the inactive state, and move apart on protonation at lower pH values[7]. This cycle converged to a stable lorazepam docking pose (Fig. 2f), and to its ranking first among the 622 decoy molecules. This strategy resembles previous ligand-guided docking[27–29], although here the binding site was unknown. In its docked geometry, lorazepam hydrogen bonds with Glu160, Arg189, Tyr244 and Tyr268, and forms non-polar contacts with Trp77, Leu101, Phe173, His269 and Leu272 (Fig. 2f).

To test the modelled lorazepam site, we mutated the Glu160, Arg189 and His269 residues lining the site (Fig. 2f and Extended Data Fig. 2e, f), and determined their roles in proton-mediated cAMP production and calcium release (Extended Data Fig. 4). The His269Phe mutant right-shifted proton concentration–response curves in both assays[7], while substitutions at Arg189 selectively abolished cAMP production. Different substitutions at Glu160 had varying effects at downstream signalling pathways—Glu160Ala left-shifted the proton concentration–response curve and reduced cAMP production, but was inactive in calcium release, while the Glu160Lys and Glu160Gln mutations had modest effects in both pathways (the mutants had little effect on expression, Extended Data Fig. 4c). These substantial and differential effects on downstream coupling support a role for these residues in the functions of GPR68, and are consistent with the modelled binding site for lorazepam.

Seeking optimized PAMs, we computationally docked 3.1 million available lead-like molecules against the putative lorazepam site in GPR68. Overall, more than 3.3 trillion complexes were calculated and scored. From among the top 0.1% of the docking-ranked molecules, 17 were purchased for testing; along with their high docking ranks, these compounds recapitulated key interactions made by lorazepam in its docked model, were chemically diverse and had high-scoring analogues (Supplementary Table 2).

Four of the docking hits increased cAMP production by about 1.5-fold over basal at pH 6.5 (Fig. 3a). Although none was as active as lorazepam, two compounds, ZINC4929116 and ZINC32587282, had hundreds of available analogues. These were docked against the GPR68 model, and 25 were chosen for testing (Fig. 3a and Supplementary Table 3). Thirteen had greater activity than lorazepam, and their pH-dependent potentiation activity clearly indicates allostery. Although dissimilar, lorazepam and ZINC67740571 dock to form many of the same interactions, with the addition of a new predicted hydrogen-bond to Glu160 from the hydroxyl of ZINC67740571 (Fig. 2f, h and Extended Data Fig. 2h).

Ogerin as a selective GPR68 PAM

Ten selected compounds were studied further in functional assays. According to the standard allosteric operational model[30], all were

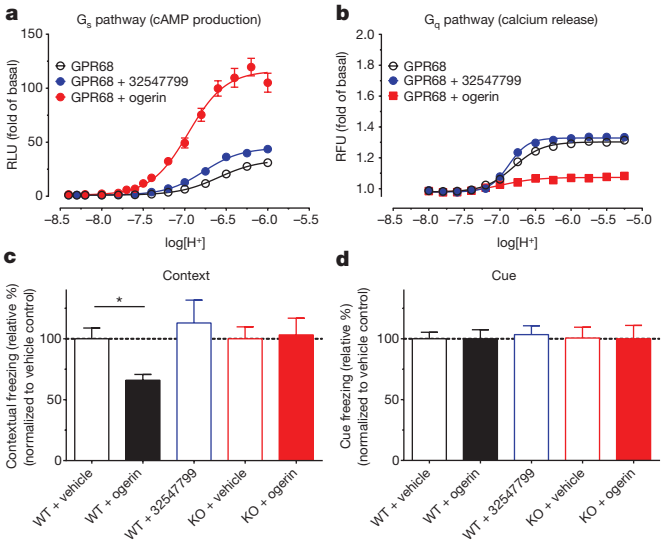


Figure 4 | Ogerin modulates signalling and memory. a, **b**, Ogerin and ZINC32547799 (10 μM) modulate proton-mediated cAMP production (**a**, $n = 4$) and calcium mobilization (**b**, $n = 5$). Data in **a** and **b** are mean ± s.e.m. RFU, relative fluorescence units. **c**, **d**, Ogerin but not its isomer (ZINC32547799) decreased contextual memory retrieval in wild-type (WT; $n = 7$) but not GPR68-knockout (KO; $n = 8$) C57BL/6J male mice (**c**, $F_{(1,27)} = 4.71$, $P < 0.05$ for drug × genotype effect, $P < 0.05$ for ogerin at wild-type mice, two-way analysis of variance (ANOVA), Bonferroni's post-hoc test); both had no effect on cued memory retrieval in either wild-type ($n = 6$) or knockout ($n = 7$) C57BL/6J male mice (**d**). Results (**c**, **d**) were normalized to vehicle control; see also Extended Data Fig. 8d–i.

GPR68 PAMs, lacking intrinsic activity but increasing agonist potency (α-factor) for cAMP production by 1.9–8.2-fold, and increasing efficacy (β-factor) by 1.1–5.6-fold (Supplementary Table 5). It is this ability to shift concentration–response curves leftward and upward (Extended Data Fig. 4b) that are the key characteristics of a PAM. ZINC67740571 had a much higher allosteric effect than lorazepam (Fig. 3b versus Fig. 1d, and Supplementary Table 8); we denoted it 'ogerin' (for OGR1 ligand).

Ogerin and ZINC32547799 are close analogues (Fig. 3a), but each had distinct functional activities (Fig. 4a and Extended Data Fig. 4f, g) and docking poses (Fig. 2 and Extended Data Fig. 2h). Thus, the ortho-hydroxylmethyl group, which differentiates them, may have a key role in determining PAM activity, perhaps because of its ability to hydrogen-bond with Glu160, which the meta-positioned hydroxylmethyl in ZINC32547799 cannot reach. The structure-guided mutants His269Phe and Arg189Leu responded to ogerin and ZINC32547799 differently (Fig. 4a, Extended Data Fig. 4f, g and Supplementary Table 6), supporting the modelled interactions with these residues. Notably, rather than activating, ogerin inhibited proton-mediated calcium release—a pathway-specific function rescued in Arg189Leu and His269Phe (Fig. 4b, Extended Data Fig. 4h, i and Supplementary Table 7). Meanwhile, ZINC32547799 had little effect on calcium release. To determine whether fast kinetics affect the difference between cAMP measurement (under equilibrium) and calcium release (non-equilibrium), we also conducted phosphatidylinositol hydrolysis assays under equilibrium. Ogerin slightly potentiated proton activity here (Extended Data Fig. 4j, k), whereas ZINC32547799 did not. Furthermore, ogerin had minimal PAM activity at the related proton-sensing GPCRs, GPR4 and GPR65 (Extended Data Fig. 2c, d). Ogerin seems to be a functionally selective GPR68 PAM for the agonist proton.

If the ogerin–GPR68 model is relevant, we should be able to leverage it for optimization. We designed a virtual library of more than 600 ogerin analogues and docked each into the GPR68 model (Extended Data Fig. 2h, i). Thirteen high-scoring analogues were synthesized, and three were more active than ogerin (Supplementary Table 9 and Extended Data Fig. 6), including the first and seventh ranked compounds, the latter of which, **C2**, had the greatest allosteric effect, shifting the proton response threefold further to the left than does ogerin, for an α-factor of 22 (Fig. 3a–c and Supplementary Table 8). **C2** differs from ogerin by the addition of a methylene to the benzylamine side chain, which places the phenyl ring deeper into a modelled apolar pocket (Extended Data Fig. 2i). The addition of one or two further methylenes in compounds **C3** and **C4** (Fig. 3a), conversely, reduced allostery (Supplementary Table 8 and Extended Data Fig. 6f), consistent with reduced complementarity to the apolar pocket in the modelled complex.

To investigate ogerin specificity for GPR68 over unrelated targets, which might affect its usefulness as a biological probe, we first computationally screened ogerin and its analogues for off-targets using the Similarity Ensemble Approach (SEA) program[31] against a panel of 2,800 targets. These calculations revealed similarity between the GPR68 ligands and those of only three other GPCRs: the ghrelin and adenosine A_1 and A_{2A} receptors. Subsequent physical profiling against 58 GPCRs, ion channels and transporters (Extended Data Fig. 3) revealed that ogerin had moderate affinity at two GPCRs, 5-hydroxytryptamine 2B (5-HT_{2B}) and the A_{2A} receptor (Extended Data Fig. 5h, i), the latter consistent with the SEA prediction.

Intrigued by the association between the GPR68 PAMs and adenosine receptor antagonists, we computationally screened a library (http://www.tocris.com/dispprod.php?ItemId=5386#.U_s5ZMVdUrU) of 1,120 reagents and drugs against the GPR68 ligands, again using SEA. SLV320, a selective adenosine A_1 antagonist[32], was predicted to be a GPR68 PAM and confirmed by a physical screen of the full library (SLV320 $\alpha\beta = 2.8$) (Extended Data Fig. 7 and Supplementary Table 8), as was a second adenosine receptor antagonist, CGH2466 ($\alpha\beta = 2.9$), and tracazolate ($\alpha\beta = 3.4$), a GABAergic (GABA-mediated) drug that also antagonizes adenosine receptors[33]. Although CGH2466 has the lowest apparent binding constant (K_B) of any GPR68 PAM (48 nM), its allostery is much lower than that of ogerin; additionally, like SLV320 and tracazolate, CGH2466 is a potent phosphodieseterase inhibitor (Extended Data Fig. 7) and had minimal activity in the presence of Ro 20-1724. This previously unknown cross-talk among the GPR68, adenosine and GABA receptor ligands (Extended Data Fig. 7d), along with their activities at phosphodiesterases, should be considered when evaluating the pharmacology of what have been considered specific probes and drugs.

Ogerin as a GPR68 probe

Given its activity and specificity, we sought to explore the downstream signalling and *in vivo* activity of ogerin. In GPR68-expressing HEK293 cells, we found that both ogerin and lorazepam activate the protein kinase A (PKA) and mitogen-activated protein (MAP) kinase pathways (Extended Data Fig. 8a), mimicking the low pH-induced signalling observed with GPR68 receptors in human airway smooth muscle cells[19]. The activation of GPR68 in smooth muscle cells by extracellular acidification is linked to several downstream pathways and biological responses[18,19,22,34–37], which a selective allosteric modulator, such as ogerin, may help to disentangle.

To investigate effects in behaviour associated with modulation of the hippocampus, where GPR68 is highly expressed[17], we evaluated GPR68-knockout and wild-type mice in a learning and memory test, fear conditioning, in which the hippocampus has important roles (Extended Data Figs 8 and 11). In wild-type mice, ogerin attenuated contextual-based fear memory without effects on cue-based memory (Fig. 4c, d). The magnitude of these effects is comparable to those of compounds targeting other hippocampus-expressed GPCRs[38,39], and larger effects are rarely observed without surgical lesion of the hippocampus[40]. Crucially, the administration of ogerin had no effect on memory retrieval in GPR68-knockout mice (Fig. 4c, d), indicating that the *in vivo* effects of ogerin are GPR68-dependent. Furthermore, the less active ogerin isomer, ZINC32547799, had no measurable effect

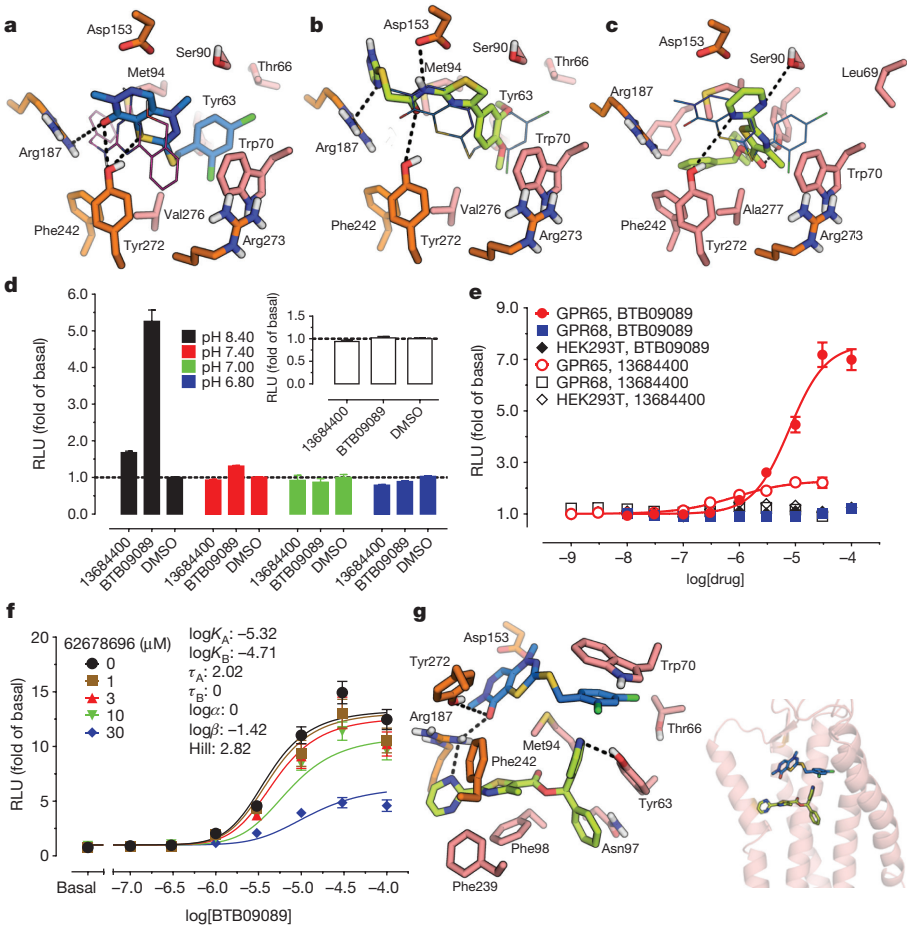


Figure 5 | Discovery of GPR65 allosteric agonist and negative allosteric modulator. a–c, Predicted interactions of BTB09089 (**a**), ZINC13684400 (**b**) and ZINC62678696 (**c**) with GPR65. Overlaid ogerin (thin magenta lines) (**a**) or BTB09089 (thin blue lines) docking poses with GPR68 or GPR65, respectively (**b**, **c**). **d**, ZINC13684400 (30 μM) displayed GPR65 allosteric agonist activity at pH 8.40 but not at lower pH or in control cells ($n = 32$ measurements). **e**, ZINC13684400 as a GPR65 agonist at pH 8.40 ($n = 3$). **f**, ZINC62678696 shifts BTB09089 curves downward at pH 8.40 ($n = 4$). K_A and K_B are the equilibrium binding affinities of the orthosteric agonist proton (A) and allosteric modulator (B), respectively. Normalized results (**d**–**f**) are mean ± s.e.m., and curves were analysed using a four-parameter logistic function (**e**) or a standard operational allosteric model (**f**). **g**, Predicted ternary complex between GPR656, ZINC62678696 and BTB09089, detailed interactions (left) and overall orientation in the GPR65 structure (right).

on learning and memory in wild-type mice (Fig. 4c, d and Extended Data Fig. 8d–i). The effects of ogerin thus support a role for GPR68 in hippocampal-associated memory.

General applicability of the approach

To explore the broader usefulness of this approach, we sought ligands for GPR65, another understudied pH-sensing receptor, which shares 37% sequence identity to GPR68. We found that a recently reported GPR65 agonist BTB09089 (ref. 41) is an allosteric agonist of GPR65 (Fig. 5d, e and Extended Data Fig. 10a). We used BTB09089 to anchor modelling of GPR65, generating 500 homology models templated on GPR68. The final docked GPR65–BTB09089 model resembles that of GPR68–ogerin, with several side-chain substitutions in the putative binding site (Fig. 5a).

We docked the same 3.1-million compounds against the GPR65 model, purchasing 45 new molecules for testing (Fig. 5a–c and Supplementary Table 10). ZINC13684400 showed agonist activity of more than twofold of basal at GPR65, with a potency of 500 nM, without measurable activity at control cells (Fig. 5e and Extended Data Fig. 9). As with BTB09089, ZINC13684400 did not potentiate proton efficacy at GPR65 (Fig. 5d), but acted as an allosteric agonist. To test the model, three residues modelled to interact with both BTB09089 and ZINC13684400, Arg187, Phe242 and Tyr272, were mutated, as was Asp153, which appears to only hydrogen-bond with ZINC13684400 (Fig. 5b). Arg187Leu, Phe242Ala and Tyr272Ala reduced the activity of both compounds (Extended Data Fig. 10f, g), whereas Asp153Ala had no effect on BTB09089 but much reduced the activity of ZINC13684400, consistent with the model. Several other docking hits inhibited GPR65 when the receptors were activated by protons or by BTB09089, including ZINC62678696 (Extended Data Fig. 10b–d). Unexpectedly, ZINC62678696 does not compete with BTB09089, as predicted, but rather acts as a BTB09089 negative allosteric modulator (Fig. 5f), suggesting that the two molecules can bind to GPR65 simultaneously (Fig. 5g).

Discussion

A combined empirical and structure-based approach discovered potent PAMs at the understudied receptor GPR68, and an allosteric agonist and negative allosteric modulators for the understudied GPR65. This supports the usefulness of the approach for illuminating the 'dark matter' of the GPCRs—the 38% of non-olfactory GPCR targets whose ligands and function are understudied or unknown[1]. Whereas truly high-throughput screens are impractical for targets of unknown function, lower-throughput screens are often feasible. Although the hits from such a screen may be unsuitable as probes, they can anchor computational screens for more optimized compounds. Correspondingly, we would not ordinarily expect docking to succeed against models of a target that shares only 29% sequence identity with its nearest template. By calculating several thousand models, and insisting that the relevant ones are those that prioritize active over inactive molecules,

functionally relevant models are prioritized. The new ligands that emerged are specific for the target and one is active *in vivo*, supporting their use as chemical probe for the function of GPR68.

Pharmacologically, the most unexpected observation was the activity of GPR68 in learning and memory. Previous studies in GPR68-knockout mice revealed only modest phenotypic changes[16,21,22], none in higher brain function, even though GPR68 is most highly expressed in the brain. Ogerin transiently and reversibly reduced contextual-based fear memory in wild-type but not GPR68-knockout mice, consistent with on-target activity *in vivo*. In hindsight, this is perhaps only accessible to chemical modulators, which can have PAM activities. Inhibitory genetic perturbations, such as knockouts or knockdowns, although crucial to demonstrating on-target activity through chemical genetic epistasis, cannot on their own reveal such activation-based modulation.

Deorphanizing a receptor can also illuminate its off-target roles for known drugs. The observation that lorazepam and its primary metabolite, desmethyldiazepam, are GPR68 PAMs may clarify several of the idiosyncratic effects of this widely used anxiolytic. Lorazepam, uniquely among benzodiazepines, can treat catatonia, an effect proposed to involve an unknown secondary target[42]. GPR68 may have a role in this efficacy, as both drug and metabolite reach micromolar concentrations in plasma during treatment[43].

Certain caveats bear airing. The combination of empirical and computational screens will not work for all orphan receptors. GPCRs that are poorly expressed or non-functional in yeast or transfected cells will be problematic, and some orphans will simply not recognize any of the molecules screened in the small empirical libraries. Also, some orphans will bear too little similarity to templates of known structure to support accurate modelling. Even those that do work will demand cycles of testing and optimization, which was crucial for both GPR65 and GPR68.

These cautions should not obscure the key observations from this study—that combining empirical and structure-based screening led to a probe molecule that reveals some of the functions of GPR68. The finding that ogerin potentiates GPR68 activation and downstream MAP kinase pathways, and previous observations that the receptor mediates airway inflammation, enables campaigns for GPR68 PAMs that may regulate respiratory inflammatory responses. Uniquely as PAMs, these compounds would have fidelity to the natural spatial and temporal activation of GPR68. Correspondingly, the role of GPR68 in anxiety offers a new route to treating this condition and related CNS disorders, an area in need of new therapeutic modalities[44]. Methodologically, this approach may have broad application to illuminating the function of the dark matter of the genome, that still large area of pharmacology in which targets are known, but function is hidden.

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

Received 2 January; accepted 4 September 2015.
Published online 9 November 2015.

1. Roth, B. L. & Kroeze, W. K. Integrated approaches for genome-wide interrogation of the druggable non-olfactory G protein-coupled receptor superfamily. *J. Biol. Chem.* **290,** 19471–19477 (2015).
2. Davenport, A. P. *et al.* International Union of Basic and Clinical Pharmacology. LXXXVIII. G protein-coupled receptor list: recommendations for new pairings with cognate ligands. *Pharmacol. Rev.* **65,** 967–986 (2013).
3. Chung, S., Funakoshi, T. & Civelli, O. Orphan GPCR research. *Br. J. Pharmacol.* **153** (suppl. 1), S339–S346 (2008).
4. Knapp, S. *et al.* A public–private partnership to unlock the untargeted kinome. *Nature Chem. Biol.* **9,** 3–6 (2013).
5. Ferguson, F. M. *et al.* Targeting low-druggability bromodomains: fragment based screening and inhibitor design against the BAZ2B bromodomain. *J. Med. Chem.* **56,** 10183–10187 (2013).
6. Leung, D., Hardouin, C., Boger, D. L. & Cravatt, B. F. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. *Nature Biotechnol.* **21,** 687–691 (2003).
7. Ludwig, M. G. *et al.* Proton-sensing G-protein-coupled receptors. *Nature* **425,** 93–98 (2003).
8. Mogi, C. *et al.* Sphingosylphosphorylcholine antagonizes proton-sensing ovarian cancer G-protein-coupled receptor 1 (OGR1)-mediated inositol phosphate production and cAMP accumulation. *J. Pharmacol. Sci.* **99,** 160–167 (2005).
9. Li, J. *et al.* Ovarian cancer G protein coupled receptor 1 suppresses cell migration of MCF7 breast cancer cells via a Gα12/13-Rho-Rac1 pathway. *J. Mol. Signal.* **8,** 6 (2013).
10. Singh, L. S. *et al.* Ovarian cancer G protein-coupled receptor 1, a new metastasis suppressor gene in prostate cancer. *J. Natl. Cancer Inst.* **99,** 1313–1327 (2007).
11. Schneider, J. W. *et al.* Coupling hippocampal neurogenesis to brain pH through proneurogenic small molecules that regulate proton sensing G protein-coupled receptors. *ACS Chem. Neurosci.* **3,** 557–568 (2012).
12. Frick, K. K., Krieger, N. S., Nehrke, K. & Bushinsky, D. A. Metabolic acidosis increases intracellular calcium in bone cells through activation of the proton receptor OGR1. *J. Bone Miner. Res.* **24,** 305–313 (2009).
13. Komarova, S. V., Pereverzev, A., Shum, J. W., Sims, S. M. & Dixon, S. J. Convergent signaling by acidosis and receptor activator of NF-κB ligand (RANKL) on the calcium/calcineurin/NFAT pathway in osteoclasts. *Proc. Natl Acad. Sci. USA* **102,** 2643–2648 (2005).
14. Yang, M. *et al.* Expression of and role for ovarian cancer G-protein-coupled receptor 1 (OGR1) during osteoclastogenesis. *J. Biol. Chem.* **281,** 23598–23605 (2006).
15. Russell, J. L. *et al.* Regulated expression of pH sensing G protein-coupled receptor-68 identified through chemical biology defines a new drug target for ischemic heart disease. *ACS Chem. Biol.* **7,** 1077–1083 (2012).
16. Mohebbi, N. *et al.* The proton-activated G protein coupled receptor OGR1 acutely regulates the activity of epithelial proton transport proteins. *Cell. Physiol. Biochem.* **29,** 313–324 (2012).
17. Regard, J. B., Sato, I. T. & Coughlin, S. R. Anatomical profiling of G protein-coupled receptor expression. *Cell* **135,** 561–571 (2008).
18. Chen, Y. J., Huang, C. W., Lin, C. S., Chang, W. H. & Sun, W. H. Expression and function of proton-sensing G-protein-coupled receptors in inflammatory pain. *Mol. Pain* **5,** 39 (2009).
19. Saxena, H. *et al.* The GPCR OGR1 (GPR68) mediates diverse signalling and contraction of airway smooth muscle in response to small reductions in extracellular pH. *Br. J. Pharmacol.* **166,** 981–990 (2012).
20. Wang, J., Sun, Y., Tomura, H. & Okajima, F. Ovarian cancer G-protein-coupled receptor 1 induces the expression of the pain mediator prostaglandin E2 in response to an acidic extracellular environment in human osteoblast-like cells. *Int. J. Biochem. Cell Biol.* **44,** 1937–1941 (2012).
21. Li, H. *et al.* Abnormalities in osteoclastogenesis and decreased tumorigenesis in mice deficient for ovarian cancer G protein-coupled receptor 1. *PLoS ONE* **4,** e5705 (2009).
22. Aoki, H. *et al.* Proton-sensing ovarian cancer g protein-coupled receptor 1 on dendritic cells is required for airway responses in a murine asthma model. *PLoS ONE* **8,** e79985 (2013).
23. Mogi, C., Nakakura, T. & Okajima, F. Role of extracellular proton-sensing OGR1 in regulation of insulin secretion and pancreatic β-cell functions. *Endocr. J.* **61,** 101–110 (2013).
24. Okajima, F. Regulation of inflammation by extracellular acidification and proton-sensing GPCRs. *Cell. Signal.* **25,** 2263–2271 (2013).
25. Dong, S., Rogan, S. C. & Roth, B. L. Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs. *Nature Protocols* **5,** 561–573 (2010).
26. Mysinger, M. M. & Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50,** 1561–1573 (2010).
27. Evers, A. & Klebe, G. Ligand-supported homology modeling of G-protein-coupled receptor sites: models sufficient for successful virtual screening. *Angew. Chem. Int. Ed. Engl.* **43,** 248–251 (2004).
28. Cavasotto, C. N. *et al.* Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J. Med. Chem.* **51,** 581–588 (2008).
29. Katritch, V., Rueda, M., Lam, P. C.-H., Yeager, M. & Abagyan, R. GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins* **78,** 197–211 (2010).
30. Leach, K., Sexton, P. M. & Christopoulos, A. Allosteric GPCR modulators: taking advantage of permissive receptor pharmacology. *Trends Pharmacol. Sci.* **28,** 382–389 (2007).
31. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nature Biotechnol.* **25,** 197–206 (2007).
32. Kalk, P. *et al.* The adenosine A1 receptor antagonist SLV320 reduces myocardial fibrosis in rats with 5/6 nephrectomy without affecting blood pressure. *Br. J. Pharmacol.* **151,** 1025–1032 (2007).
33. Thompson, S.-A., Wingrove, P. B., Connelly, L., Whiting, P. J. & Wafford, K. A. Tracazolate reveals a novel type of allosteric interaction with recombinant γ-aminobutyric $acid_A$ receptors. *Mol. Pharmacol.* **61,** 861–869 (2002).
34. Tomura, H. *et al.* Prostaglandin I_2 production and cAMP accumulation in response to acidic extracellular pH through OGR1 in human aortic smooth muscle cells. *J. Biol. Chem.* **280,** 34458–34464 (2005).
35. Ichimonji, I. *et al.* Extracellular acidification stimulates IL-6 production and Ca^{2+} mobilization through proton-sensing OGR1 receptors in human airway smooth muscle cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **299,** L567–L577 (2010).

36. Liu, J. P. *et al.* Ovarian cancer G protein-coupled receptor 1-dependent and -independent vascular actions to acidic pH in human aortic smooth muscle cells. *Am. J. Physiol. Heart Circ. Physiol.* **299,** H731–H742 (2010).
37. Matsuzaki, S. *et al.* Extracellular acidification induces connective tissue growth factor production through proton-sensing receptor OGR1 in human airway smooth muscle cells. *Biochem. Biophys. Res. Commun.* **413,** 499–503 (2011).
38. Gravius, A., Barberi, C., Schäfer, D., Schmidt, W. J. & Danysz, W. The role of group I metabotropic glutamate receptors in acquisition and expression of contextual and auditory fear conditioning in rats – a comparison. *Neuropharmacology* **51,** 1146–1155 (2006).
39. Daumas, S. *et al.* Transient activation of the CA3 Kappa opioid system in the dorsal hippocampus modulates complex memory processing in mice. *Neurobiol. Learn. Mem.* **88,** 94–103 (2007).
40. Phillips, R. G. & LeDoux, J. E. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav. Neurosci.* **106,** 274–285 (1992).
41. Onozawa, Y. *et al.* Activation of T cell death-associated gene 8 regulates the cytokine production of T cells and macrophages *in vitro*. *Eur. J. Pharmacol.* **683,** 325–331 (2012).
42. Pompéia, S., Manzano, G. M., Tufik, S. & Bueno, O. F. What makes lorazepam different from other benzodiazepines? *J. Physiol. (Lond.)* **569,** 709 (2005).
43. Greenblatt, D. J. *et al.* Clinical pharmacokinetics of lorazepam. I. Absorption and disposition of oral 14C-lorazepam. *Clin. Pharmacol. Ther.* **20,** 329–341 (1976).
44. Schoepp, D. D. Where will new neuroscience therapies come from? *Nature Rev. Drug Discov.* **10,** 715–716 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by National Institutes of Health (NIH) grants U01104974 (B.L.R., B.K.S. and W.K.K.), R01 DA017204 (B.L.R. and W.K.K.) and the National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP) (X.-P.H., H.Z., M.S.F., W.K.K., T.J.M., A.J. and B.L.R.), the Michael Hooker Chair for Protein Therapeutics and Translational Proteomics to B.L.R.; Genentech Foundation Predoctoral Fellowship (J.K.); NIH grants GM59957 and GM71896 (B.K.S.) and the Structural Genomics Consortium (B.K.S.); grant P01 HL114471 (R.B.P. and D.A.D.); NICHD grant U54 HD079124 (M.S.M., K.A.S., V.N.); NIH grant U19MH082441 (B.L.R., J.J. and X.C.). We thank Mark Pausch (Merck & Co.) for providing us G_s- and G_q-yeast strains for yeast screening assays.

Author Contributions X.-P.H. subcloned GPR68 for yeast screening, made GPR68 and GPR65 mutants, designed, carried out cell-based screening assays, analysed results, and wrote the paper. J.K. designed and developed homology models, carried out docking screens, analysed results, and wrote the paper. W.K.K. set up and performed yeast screening assays, analysed results, and wrote the paper. H.Z. and M.S.F. designed, performed *in vivo* fear-conditioning studies, analysed results, and wrote the paper. M.S.F. and B.L.R. dubbed ZINC67740571 'ogerin'. B.H.K. created the GPR68-knockout mice. S.S.M., K.A.S. and V.N. carried out initial phenotypic characterization, analysed results, and wrote the paper. X.C. and J.J. synthesized ZINC32547799, ZINC67740571 (ogerin) and ogerin analogues (compounds 33548–33561, **C3** and **C4**) for functional assays and *in vivo* studies, and wrote the paper. T.J.M. carried out radioligand binding assays. A.J. prepared drug plates and plasmids for initial screening. R.B.P. and D.A.D. designed and carried out anti-haemagglutinin immunoblot assays, analysed results, and wrote the paper. S.W. designed primers, prepared Flag-tagged GPR68 wild-type and mutant plasmids, performed anti-Flag western blot assays, and analysed results. T.K. analysed results and wrote the paper. B.L.R. and B.K.S. coordinated and supervised the project, and with the other authors wrote the paper.

Additional Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.L.R. (bryan_roth@med.unc.edu) or B.K.S. (bshoichet@gmail.com).

METHODS

Chemicals, reagents and cells lines. Chemicals and reagents used in this study, if not specified otherwise, were purchased from commercial sources (Sigma, Tocris, Fisher Scientific, or specified in Supplementary Tables 2 and 3 of chemical structures) or synthesized as outlined in the Supplementary Information. HEK293 (ATCC CRL-1573; 60113019; certified mycoplasma free and authentic by ATCC) and HEK293-T (HEK293T; ATCC CRL-11268; 59587035; certified mycoplasma free and authentic by ATCC) cells were from the ATCC. Cells were also validated by analysis of short tandom repeat (STR) DNA profiles and these profiles showed 100% match at the STR database from ATCC. Ogerin and its inactive analogue ZINC32547799 are available for use as chemical probes from Sigma-Aldrich (ogerin: SML1482, ZINC32547799: SML1483).

Homology modelling. The alignment for the construction of the GPR68 models was generated using PROMALS3D, and homology models were built with MODELLER-9v8 (ref. 45), using the crystal structure of the chemokine CXCR4 receptor (PDB code 3ODU) as the template (Extended Data Fig. 2f). This alignment was also used to generate 500 models of GPR65 directly from the final GPR68 model. The initial alignment included both human and mouse sequences of GPR68, as well as those of its closest homologue, GPR4. These were aligned against the whole human C-X-C chemokine receptor family. The alignment was manually edited to: remove the amino and carboxy termini that extended past the template structure, remove the engineered T4 lysozyme, and create different alignments of the flexible and non-conserved second extracellular loop (the final result is given in the provided alignment, Extended Data Fig. 2f). A total of 407 models were built directly based on the CXCR4 crystal structure, using MODELER-9v8 (ref. 45), while five more were built from each of 580 elastic network models (ENMs), produced by the program 3K-ENM[46], for a total of 3,307 models built during each iterative round of model refinement. Models with constraints between pairs of extracellular His residues (His17–His169, His17–His269, His17–His84 and His84–His169) to mimic the inactive state of the protein were generated by enforcing a distance constraint of 2.7 Å between the imidazole nitrogens, with a standard deviation of 0.1 Å. Confirmed active compounds and analogues using CXCR4-based model had neither agonist nor antagonist activity at CXCR4 receptors (Extended Data Fig. 5j, k).

Model evaluation. Before docking, the second extracellular loop (EL2), between residues 161–177, was removed from each GPR68 model. Models were ranked on the basis of prioritizing active benzodiazepines (lorazepam and desmethyldiazepam) over the rest of the inactive NCC library that was used in the yeast screen, as well as over property-matched decoys. In addition, the docked pose of lorazepam had to form a hydrogen bond from its N–H group to a polar side chain in GPR68. Five different sites were sampled for possible lorazepam binding, based on the locations of the co-crystallized CXCR4 small molecule antagonist 1T1t (in PDB code 3ODU), cyclic peptide CVX15 (in PDB code 3OE0), and the positions of the biogenic amines crystallized with the β_2-adrenergic receptor (PDB code 2RH1) and the dopamine D_3 receptor (PDB code 3PBL). The entire NCC library was docked to each of the five sub-sites for several rounds of iterative binding site refinement. In each round, the top-ranked models were examined for a binding pose that made hydrophobic and electrostatic interactions with the receptor, including the key N–H hydrogen bond. Residues within 6 Å of the lorazepam pose were minimized around the docked ligand with PLOP[47]. The NCC library was then re-docked into this optimized binding site for each model. This refinement continued for several cycles until the top-ranked models all converged to the same lorazepam pose. Once the final model was chosen, we built the EL2 back onto the receptor using MODELLER-9v8 (ref. 45) and optimized 1,000 different EL2 conformations around the lorazepam pose with PLOP. Finally, we docked the NCC library back into these 1,000 different EL2-GPR68 structures, and chose a final model that retained the previous pose and prioritized the active over the inactive compounds. The GPR65 model was generated similarly, using the pose of BTB09089 as the primary selection criterion, although in this case the EL2 was always present. To determine the ternary complex model of ZINC62678696 and BTB09089, ZINC62678696 was docked to the putative binding site in the GPR65 model with BTB09089 present. Then, both ligands were minimized with PLOP. Next, the side chains of the GPR65 binding pocket were allowed to relax, and, finally, BTB09089 and ZINC62678696 were simultaneously minimized again with PLOP. Structural models (PDB files) of characteristic GPR68-modelled complexes (with ogerin or lorazepam) and GPR65-modelled complexes (with BTB09089 or BTB09089 and ZINC62678696) are shown in the Supplementary Data.

Virtual screens. We used DOCK 3.6 to screen the ZINC database (Results). The flexible ligand sampling algorithm in DOCK 3.6 superimposes atoms of the docked molecule onto binding site matching spheres, which represent favourable positions for individual ligand atoms. Forty-five matching spheres were used, using the previous refinement round's pose of lorazepam. The degree of ligand sampling is determined by the bin size, bin size overlap and distance tolerance, set at 0.4 Å, 0.1 Å and 1.5 Å, respectively, for both the matching spheres and the docked molecules. The complementarity of each ligand pose was scored as the sum of the receptor–ligand electrostatic and van der Waals' interaction energies, and corrected for context-dependent ligand desolvation. Partial charges from the united-atom AMBER force field were used for all receptor atoms; ligand charges and initial solvation energies were calculated using AMSOL[48,49] (http://comp.chem.umn.edu/amsol/). The best-scoring conformation of each docked molecule was then subjected to 100 steps of rigid-body minimization.

Selection of potential ligands for testing. We docked the approximately 3.1 million commercially available molecules of the lead-like subset of the ZINC database to the final GPR68 and GPR65 models. The full hit list was automatically filtered to remove molecules that possess high-internal-energy, non-physical conformations, which are not well-modelled by our scoring function. The reported rankings reflect this filtering. From the top 0.1% (~3,000 molecules) of the docked ranking list, 17 compounds were chosen for testing, based on complementarity to the binding site and presence of predicted electrostatic interactions with Glu160, Arg189, Tyr244, Tyr268 and His269, mimicking those predicted for lorazepam. For GPR65, compounds were chosen based on complementarity to the binding site and similarity to the predicted binding pose of BTB09089, modelled to interact with Asp153, Arg187 and Tyr272, and by aromatic stacking with Trp70.

***In silico* lead profiling.** To examine specificity and to discover other potential GPCR targets for the newly discovered GPR68 PAMs, we used the SEA program[31,50], which compares individual ligands and sets of ligands to the ligand sets for multiple targets; two targets are related, or a particular ligand is predicted to modulate a target, if the sets of ligands are related to one another. Here, the query set was all of the new GPR68 PAMs, which was screened against either the 2,512 ligand–target set with activity of 10 μM or better from the ChEMBL12 database[51], or against the Tocris Mini library.

Receptor constructs and yeast growth assays. Twenty-four human GPCR plasmids (GPR1, GPR4, GPR15, GPR31, GPR39, GPR41, GR43, GPR45, GPR55, GPR57, GPR58, GPR62, GPR65, GPR68, GPR83, GPR84, GPR87, GPR88, GPR123, GPR132, GPR133, GPR157, GPR161 and ADCYAP1R1) were obtained from http://cdna.org, subcloned into the multiple cloning site of the yeast high copy number plasmid p426GPD (ref. 52) and were confirmed by full-length sequencing (Eton Bioscience). The yeast strains used were provided by M. Pausch (Merck) and have been previously described[53] and used by us[25,54] ; MPY578t (G_i yeast), MPY578q5 (G_q yeast) and MPY578s5 (G_s yeast) express chimaeric G proteins in which the last five amino acids of the yeast Gα protein are replaced with their mammalian G_i, G_q or G_s homologues, respectively. These strains contain the *HIS3* gene under the control of the *FUS1* promoter. GPCR transformants in yeast were selected and maintained on synthetic defined (SD) media lacking uracil (Clontech). $GPR68_q$ indicates the GPR68 paired with G_q yeast; while $GPR4_s$ indicates GPR4 paired with G_s yeast, and similarly for the other GPCRs. The yeast screening assays were carried out as described previously[25]. Assays were set up in 96-well flat-bottom clear assay plates that contained 50 μl of test compound at 40 μM (final concentration of 10 μM, in triplicate) diluted in SD-His-Ura medium (Clontech), 50 μl of 3-amino-1,2,4-triazole (3-AT) at 4× concentration diluted in SD-His-Ura medium (pH 5.4), and 100 μl of yeast cell suspension diluted in SD-His-Ura medium to a final $A_{600\,nm}$ of 0.02. Growth was at 30 °C for 2–5 days. Before measurement of cell growth, cells were re-suspended by repeated gentle pipetting to ensure uniform suspension of cells. Cell growth was measured by absorbance at 600 nm in a microplate reader (POLARstar Omega, BMG Biotech). After culling of data from obviously contaminated wells, the $A_{600\,nm}$ values of each individual well were adjusted as follows: 100× ($A_{600\,nm}$ of test well − $A_{600\,nm}$ of plate median value) to give percentage growth stimulation (positive values), or percentage growth inhibition (negative value) in the form of mean ± s.e.m. of three wells.

To measure and control constitutive activity or leaky *HIS* expression, each receptor–yeast combination was plated as above in the absence of ligand over a range of concentrations of 3-AT. Concentrations of 3-AT that showed moderate yeast growth (that is, *A* values of 0.2–0.6) after 2 days at 30 °C were used in assays for drug screening. To measure concentration-dependent activity, various concentrations of cognate ligands diluted in SD-His-Ura medium were incubated with transformed yeast and appropriate concentrations of 3-AT for 2 days at 30 °C.

Site-directed mutagenesis. The GPR68 plasmid was obtained from http://cdna.org. Mutation of Glu160Ala, Glu160Lys, Glu160Gln, Arg189Leu, Arg189Met and His269Phe in the GPR68 and mutation of Asp153Ala, Arg187Leu, Phe242Ala and Tyr272Ala in the GPR65 were introduced with Agilent's QuikChange II site-directed mutagenesis kit and confirmed by sequencing. To tag the receptors for comparing receptor expression levels with immunoblotting, Flag epitope tag was inserted at the C terminus of the GPR68 wild-type and mutant receptors, also using the QuikChange II site-directed mutagenesis kit. Insertion was confirmed by sequencing.

Split-luciferase based cAMP reporter assays with proton receptors. GPR4, GPR65 and GPR68 plasmids were obtained from http://cdna.org. GPR68 mutations were made and confirmed as above. Receptor-mediated G_s activation was measured using a split-luciferase reporter assay (GloSensor cAMP assay, Promega). In brief, HEK293T cells were transiently co-transfected with receptor DNA and the GloSensor cAMP reporter plasmid (GloSensor 7A). Transfected cells were plated in poly-L-Lys-coated 384-well white clear bottom cell culture plates in DMEM supplemented with 1% dialysed FBS at a density of 15,000 cells per well in a total volume of 40 µl for a minimum of 6 h. Before assays, culture medium was removed and cells were incubated with luciferin (4 mM prepared in drug buffer, pH 8.4) for 90 min at 37 °C. The drug buffer was made with 1× HBSS supplemented with 10 mM HEPES and 10 mM MES modified from[19]. TAPS was added to accommodate higher pH values for some assays; no difference was observed between different buffers under the same pH conditions. Cells plated at pH 8.4 for 6 h generated the same H^+ concentration–response curves as those plated at pH 7.4. To make individual pH solutions, the pH was adjusted with NaOH and measured at room temperature with a pH 211 Microprocessor pH meter (Hanna Instruments). To measure modulator activity under different pH conditions, modulator was mixed with pH solutions before adding to cells. To achieve the goal that drug solutions were delivered at the correct pH values, luciferin solution was removed from cell plates before addition of drug solutions at predetermined pH values. To improve solubility for some hydrophobic compounds, 1 mg ml^{-1} BSA was added to drug solutions, and it had no effect on H^+ concentration–response curves. For G_s protein activity (cAMP production), the cell plate was usually incubated at room temperature for 20 min before being counted in a luminescence counter. Results were analysed using GraphPad Prism.

Allosteric operational model and data analysis. To estimate allosteric parameters, results were fitted to the allosteric operational model[30,55] as shown in the following equation:

$$\text{Response} = \text{basal} + (E_{max} - \text{basal}) \times \frac{(\tau_A[A](K_B + \alpha\beta[B]))^n}{([A]K_B + K_AK_B + K_A[B] + \alpha[A][B])^n + (\tau_A[A](K_B + \alpha\beta[B]))^n}$$

In which:

(1) Response is the measured activity in the form of RLUs for measurement of cAMP production. If the results were normalized, the 'response' is RLU in fold of basal (with buffer control as basal).

(2) E_{max} is a system parameter, representing the maximal possible response of the system, and this value was normally constrained to the maximal reading of the corresponding experiment.

(3) Basal is the baseline in the absence of test ligand, and is constrained to the baseline of the corresponding experiment. If results were normalized to fold of basal, the 'basal' was usually 1.0.

(4) [A] and [B] represent concentrations of the orthosteric and allosteric ligands, respectively. In the case of GPR68, A is proton.

(5) K_A and K_B are the equilibrium dissociation constants of the orthosteric agonist proton (A) and allosteric modulator (B), respectively. To facilitate curve-fitting with the model, K_A is usually fixed to the binding affinity determined from traditional radioligand binding assays under the assumption that the experimentally derived binding affinity is not significantly different from the functional affinity under the condition for corresponding functional assay. Since proton binding affinity is not a measurable parameter in this assay system, the proton K_A is therefore constrained to the corresponding proton potency (EC_{50}, the proton concentration for half-maximal response) value in the absence of the allosteric ligand, under the assumption that the proton potency is not significantly different from its binding affinity when the cAMP production assay is carried out. Since protons are present at relevant concentrations at physiological pH values, for a proton receptor K_B is largely a fitting parameter without a clear physical meaning.

(6) The term τ_A is the orthosteric agonist proton efficacy parameter. Since allosteric modulators in this study showed no agonist activity, the allosteric modulator efficacy τ_B is therefore 0 and not included in the function.

(7) The term n is the slope factor linking receptor occupancy to response. Steep slopes in this study indicated high cooperativity between proton binding and receptor activation, probably reflecting the fact that the proton receptors operate within a narrow physiological pH range.

(8) The allosteric parameter α defines the mutual effect between the orthosteric agonist A and the allosteric modulator B ($\alpha > 1$ for increased affinity and $\alpha < 1$ for reduced affinity); while β defines the allosteric effect on agonist efficacy ($\beta > 1$ for increased efficacy and $\beta < 1$ for reduced efficacy).

With K_A, basal and E_{max} constrained to their corresponding values, the parameters K_B, τ_A, α, β and n are globally shared fitting parameters for a family of proton concentration–response curves in the absence and presence of increasing concentrations of a test allosteric modulator. With the above settings, most curves could be easily fitted to generate reasonable parameters. If Prism could not fit the curves, but generated 'ambiguous fitting' results, the α value was then manually constrained to an initial fitting value and systematically changed with small increments or decrements until the highest stable high affinity value (K_B) was reached. For GPR65 and GPR68, K_B represents the allosteric binding affinity in the absence of protons, which is unmeasurable and thus has little physical meaning. The value $K_B/(1+\alpha)$ represents the binding affinity of an allosteric ligand in the presence of protons, which could be estimated experimentally. For convenience, we call $K_B/(1+\alpha)$ the 'Biochemical binding affinity, K_{bB}' (Supplementary Table 8) for an allosteric ligand in the presence of an orthosteric agonist (in this case, H^+).

Calcium mobilization assays. HEK293T cells were transfected and plated into poly-L-Lys-coated 384-well black clear bottom cell culture plates in DMEM supplemented with 1% dialysed FBS, at a density of 15,000 cells in 40 µl per well for overnight. Before the assay, medium was removed and cells were loaded with Fluo-4 Direct calcium dye (Invitrogen) for 60 min at 37 °C in a 5% CO_2 atmosphere. The calcium dye was prepared in drug buffer supplemented with 2.5 mM probenecid, pH 8.0. Proton solutions were made with 1× HBSS, 7 mM HEPES, 7 mM HEPPS and 7 mM MES, and pH was adjusted with NaOH. Drug additions and fluorescence intensity measurement were carried out in a $FLIPR^{TETRA}$, which was programmed to add drug solutions to cells while recording fluorescence intensity. To measure proton concentration–responses, 10 µl of pH pre-determined solutions were added to each well (with 20 µl calcium dye) while fluorescence intensity was recorded during and after addition for 4 min (one reading per second). The addition procedure was configured in such a way (30 µl per second at height of 10 µl above cells) that local proton concentrations for cells were essentially the same as in the pH working solutions at the moment of addition. Fluorescence intensities reached peak values within 30 s after drug addition. To determine the effects of modulators on proton responses, the protocol was modified slightly. In brief, cells were loaded with calcium dye as above, but only at 15 µl per well. The $FLIPR^{TETRA}$ was programmed to first add 5 µl of 4× test compound (final concentration of 10 µM before addition of 10 µl of pH solutions) prepared with the same drug buffer at pH 8.0 (buffer alone served as a control). After a total of 10 min of reading and incubation, 10 µl of the pH solutions were added and the fluorescence intensity was recorded exactly the same way as above. Results (fluorescence intensity in fold of basal) were exported and analysed in GraphPad Prism. For calcium mobilization assays with $5\text{-}HT_{2B}$ receptors, HEK293 cells stably expressing human $5\text{-}HT_{2B}$ receptors were used instead of transiently-transfected cells. Cells were set up and tested in the same way as above, with 5-HT serving as an agonist control (3 pM–30 µM), and with 1 nM 5-HT being used in the second addition to determine the antagonist activity of ogerin.

Phosphatidylinositol hydrolysis assay. HEK293T cells were transfected for 24 h and plated in poly-L-Lys-coated 96-well black clear bottom cell culture plates in DMEM supplemented with 10% FBS, at a density of 60,000 cells in 100 µl per well. After 5 h, cells were washed with inositol-free DMEM once and labelled with ^{3}H-inositol (1 µCi per well, PerkinElmer) in inositol-free DMEM supplemented with 5% dialysed FBS overnight. On the assay day, labelling medium was removed and cells were washed once with assay buffer (1× HBSS, 10 mM HEPES, 10 mM MES and 20 mM LiCl, pH 8.4). To measure drug concentration responses, then cells were then incubated with drug solutions at pH 8.4 for 20 min. To measure proton concentration responses, the assay buffer was pre-adjusted to desired pH values and supplemented with 20 mM LiCl. To measure the effect of ogerin or its isomer ZINC32547799 on proton concentration–response curves, pH solutions were supplemented with 20 mM LiCl and 10 µM ogerin or ZINC32547799. The premixed drug solutions were added to cells for 20 min. At the end of incubation, drug solutions were removed and 40 µl per well of 50 mM ice-cold formic acid was added. After incubation at 4 °C for 30 min, the acid extracts were transferred to polyethylene terephthalate 96-well sample plates (1450-401, Perkin Elmer) and mixed with 75 µl (200 µg) YSi RNA binding beads (RPNQ0013, Perkin Elmer). The plate was sealed and further incubated at 4 °C for 30 min before being counted on a TriLux MicroBeta counter. Results (c.p.m. per well) were analysed using Graphpad Prism.

Functional assays with A_{2A} and CXCR4 receptors. Functional assays with A_{2A} adenosine and CXCR4 chemokine receptors were carried out using a slightly different protocol from that previously described for G_s (above) and G_i receptors[56]. Specifically, HEK293T cells were transfected and plated using regular DMEM supplemented with 1% dialysed FBS. Before assays, culture medium was removed, and cells were incubated with 20 µl drug solution (prepared in drug buffer 20 mM HEPES, 1× HBSS, pH 7.4) for 15 min at room temperature. To measure

agonist activity, 5 µl of 5× luciferin solution (4 mM final concentration) for A_{2A} (G_s-coupled GPCRs) or a mixture of luciferin and isoproterenol at a final concentration of 200 nM for CXCR4 (G_i-coupled GPCRs) was added and cells were incubated for another 20 min. To measure antagonist activity, test compound was added first for 10 min before a reference agonist at a final of EC_{80} concentration for another 10 min, and then followed by addition of luciferin for A_{2A} or a mixture of luciferin and isoproterenol for CXCR4 as above. Luminescence was measured in a luminescence counter. Results were analysed in GraphPad Prism.

Radioligand binding assays. Radioligand binding assays with selected CNS targets were carried out as described[56,57] and as detailed in the PDSP protocol book available online (http://pdsp.med.unc.edu/pdspw/binding.php). In brief, receptor membrane preparations were made from either animal brain tissues, or stable cell lines, or transiently transfected HEK293T cells. Receptor expression levels and radioligand binding affinities were determined with saturation binding assays. Competition binding assays were performed with membrane aliquots and a fixed concentration of radioligand in 96-well plates in a final volume of 125 µl. Reactions were incubated in the dark and at room temperature (22 °C), and terminated by vacuum filtration onto 96-well formatted GF/B filters. Radioactivity on the filters was counted in a beta counter. Results were analysed in GraphPad Prism.

Anti-HA immunoblots. HEK293 cells were transfected with either pcDNA3 vector containing a haemagglutinin (HA) cassette within the multiple cloning site, or pcDNA3HA-GPR68 encoding human GPR68 with an N-terminal HA tag. Stable lines were generated by selection with 250 µg ml^{-1} G418, with >90% of cells expressing HA after 2 weeks as assessed by immunocytochemistry (not shown). Cells were plated into 12-well plates, grown to confluence, and media switched to Hams-F12 media, with pH adjusted to pH 8.0 or 7.4, for 1 h. Cells were then stimulated with vehicle, 50 µM ogerin, or 50 µM lorazepam for 10 min. Lysates were collected and subjected to immunoblotting, with blots probed using primary antibodies against HA (Sigma cat H3663), total vasodilator-stimulated phosphoprotein (VASP, BD Biosciences, cat 610448), p-p42/p44 (Cell Signaling, cat 5726S), and β-actin (Sigma, cat A1978), and secondary antibodies (Licor, cat 926-32213 and 926–32210) conjugated with infrared fluorophores as described previously[58].

Anti-Flag immunoblots. HEK293T cells were transiently transfected in 10-cm dishes with Flag-tagged GPR68 wild-type and mutant receptors. Untransfected HEK293T cells served as a negative control. After 48 h, cells were collected, lysed and sonicated to shear chromatin before being subjected to immunoblotting. Blots were probed with monoclonal anti-Flag M2-peroxidase antibody (Sigma, A8952). Bands were quantified and normalized to GPR68 wild-type receptor (fold) for graphing.

Data analysis and reporting. Other than *in vivo* studies (below), no statistical analysis was applied to yeast- or cell-based screening assays. Sample size (number of assays for each compound or receptor) was predetermined to be in triplicate or quadruplicate for primary screening assays at a single concentration. Some samples were repeated more than the others in the primary screening assays and the number of measurements were specified as a range in corresponding figure legends. For concentration–response assays, the sample size (number of assays for each compound at selected receptors) was also predetermined to be tested for a minimum of three assays, each in triplicate or quadruplicate. Samples or receptors were tested not randomly but in an alphabetic order or numeric order according to their coded names for easy organization and were thus blinded. For each batch of assays, a control assay with isoproterenol and proton concentration–responses were included. If potency values for either isoproterenol or proton was >0.5 log unit away from established averages, assays with the batch of transfected cells were excluded. For structure–activity relationship (SAR) studies, only the assays in which all related compounds were tested side by side were included. None of the functional assays were blinded to investigators.

Generation of GPR68-knockout mice. To generate GPR68-knockout mice, a probe specific for the human *GPR68* transcript was generated by PCR amplification of a 450-base-pair (bp) segment of the coding sequence of the final exon of *GPR68* using total placental RNA. The probe was used to identify a clone from a 129 mouse genomic lambda library. The genomic insert was subcloned and a restriction map generated using a panel of enzymes. The targeting construct for the *GPR68* locus consists of a PGK-1 promoter driven neomycin resistance cassette flanked by two arms of homology with the mouse *GPR68* locus. The longer arm of homology was generated using a 7,266-bp PstI fragment extending from the last intron to the beginning of the last exon. This exon contains the entire coding sequence of the *GPR68* gene. The 1,335-bp shorter arm was generated by PCR amplification and extends from the downstream end of the long arm into the 3′ untranslated region of the gene. Homologous recombination of the targeting construct with the *GPR68* locus inserts the neomycin resistance cassette into codon 78 of the gene, thereby disrupting expression. Correctly targeted cell lines were identified by Southern blot analysis using a probe consisting of a 1,496-bp PstI fragment immediately upstream of the long arm. This probe recognizes a 14,290-bp EcoRV fragment in the endogenous locus and a 7,855-bp fragment in the targeted locus. Genotyping was carried out by PCR with three primers. The common (5′-GCAG AGGAAGCCCACGCTGATGTA-3′) and endogenous (5′-TAAACGGTAGCTGT GATTATTCAA-3′) primers generate a 516-bp PCR product from the endogenous locus, while the common and targeted (5′-AAATGCCTGCTCTTTACTGAAGG-3′) primers generate a 465-bp product from the targeted locus. The chimaeras were bred to C57BL/6J mice and pups carrying the mutant allele identified. After ten successive crosses of heterozygous animals to C57BL/6J mice, heterozygous mice were intercrossed and a congenic *Gpr68*$^{-/-}$ and C57BL/6J breeding colony established. The GPR68-knockout mice were profiled in several behavioural tests as described below in detail and results are summarized in Extended Data Fig. 11 and Supplementary Tables 11 and 12.

***In vivo* behavioural profiles of GPR68-knockout mice.** Mice were maintained and handled according to the Guide for the Care and Use of Laboratory Animals approved by the Institutional Animal Care and Use Committee of the University of North Carolina at Chapel Hill. The goal of this study was to determine whether targeted deletion of GPR68 alters behavioural function in mice.

Timeline for behavioural tests. The following tests were performed with mice at the ages shown in parentheses. Elevated plus maze test for anxiety-like behaviour (6–7 weeks); activity in an open field, accelerating rotarod (2 tests, 48 h apart) (7–8 weeks); three-chamber social approach test, activity in an open field (re-test) (8–9 weeks); marble-burying assay (9–10 weeks); acoustic startle test, buried food test for olfactory ability (10–11 weeks); visual cue test in the Morris water maze (11–12 weeks); hidden platform test for spatial learning (12–14 weeks); reversal learning in the Morris water maze (14–16 weeks); second acoustic startle test, hotplate test for thermal sensitivity (16–17 weeks).

Summary of results. Mice with deletion of GPR68 had normal performance in most of the behavioural tests. No effects of genotype were observed for body weights, activity and anxiety-like behaviour in an elevated plus maze or an open field, motor coordination, sociability, prepulse inhibition of acoustic startle responses or acquisition in the water maze. However, both male and female GPR68-knockout mice had small, significant decreases in acoustic startle responses, suggesting a reduced responsivity to environmental stimuli. Male GPR68-knockout mice also showed significant decreases in marble burying, a test for anxiety-like phenotypes. Overall, the findings indicate that GPR68 might have a role in specific domains of behaviour.

Elevated plus maze. This test is used to assess anxiety-like behaviour in rodents. The procedure is based on a natural tendency of mice to actively explore a new environment, versus a fear of being in an open area. In the present study, mice were given one 5-min trial on the plus maze, which had two walled arms (the closed arms, 20 cm in height) and two open arms. The maze was elevated 50 cm from the floor, and the arms were 30 cm long. Animals were placed on the centre section (8 × 8 cm), and allowed to freely explore the maze. Measures were taken of time on, and number of entries into, the open and closed arms. All of the experimental groups showed a strong preference for the closed arms, in comparison to the open arms, of the elevated plus maze. As shown in Supplementary Table 11, there were no significant differences between the wild-type and GPR68-knockout mice for percentage time or percentage entries on the open arms, or for total entries during the task.

Activity in an open field. Exploratory activity in a novel environment was assessed in an open field chamber (41 × 41 × 30 cm) crossed by a grid of photobeams (VersaMax system, AccuScan Instruments). Counts were taken of the number of photobeams broken during the trial in 5-min intervals, with separate measures for ambulation (total distance travelled) and rearing movements. Time spent in the centre region of the open field was measured as an index of anxiety-like behaviour. Unfortunately, an equipment malfunction led to the loss of data for 8 mice during the first activity test, conducted when mice were 7–8 weeks in age. Therefore, a second activity test was given, when mice were 8–9 weeks in age. As depicted in Extended Data Fig. 11a, b, there were no significant differences between the wild-type and GPR68-knockout mice for distance travelled, or for rearing or centre time (data not shown), during the second activity test. A significant sex × time interaction was found for the distance measure ($F_{(11,385)} = 2.68$, $P = 0.0025$), reflecting higher levels of activity in the female groups at the beginning of the session.

Accelerating rotarod test. Subjects were tested for motor coordination and learning on an accelerating rotarod (Ugo Basile). For the first test session, animals were given three trials, with 45 s between each trial. Two additional trials were given 48 h later. Revolutions per minute (rpm) was set at an initial value of 3, with a progressive increase to a maximum of 30 rpm across five minutes (the maximum trial length). Measures were taken for latency to fall from the top of the rotating barrel. As shown in Extended Data Fig. 11c, d, deletion of GPR68 did not lead to deficits in motor coordination on the rotarod. In fact, during the first three

acquisition trials, there was a non-significant trend for enhanced performance in the male knockout group (repeated-measures ANOVA, genotype × sex interaction, $F_{(1,35)} = 3.58$, $P = 0.0668$).

Marble-burying assay. This procedure is used to evaluate anxiety-like behaviour and repetitive responses. Mice were tested in a Plexiglas cage located in a sound-attenuating chamber with ceiling light and fan. The cage contained 5 cm of corncob bedding, with 20 black glass marbles (14 mm diameter) arranged in an equidistant 5 × 4 grid on top of the bedding. Animals were given access to the marbles for 30 min. Measures were taken of the number of buried marbles (two-thirds of the marble covered by the bedding). A two-way ANOVA indicated a significant genotype × sex interaction ($F_{(1,35)} = 7.37$, $P = 0.0102$) (Supplementary Table 11). Post-hoc comparisons revealed that the male GPR68-knockout mice buried significantly fewer marbles than both male wild-type mice and female knockout mice in this task.

Buried food test for olfactory function. Several days before the olfactory test, an unfamiliar food (Froot Loops, Kellogg Co.) was placed overnight in the home cages of the mice. Observations of consumption were taken to ensure that the novel food was palatable. Sixteen to twenty hours before the test, all food was removed from the home cage. On the day of the test, each mouse was placed in a large, clean tub cage (46 × 23.5 × 20 cm (width, length, height)), containing paper chip bedding (3-cm deep), and allowed to explore for 5 min. The animal was removed from the cage, and one Froot Loop was buried in the cage bedding. The animal was then returned to the cage and given fifteen minutes to locate the buried food. Measures were taken of latency to find the food reward. As shown in Supplementary Table 11, there were no significant differences between the groups in latency to find the buried food.

Hotplate test for thermal sensitivity. Individual mice were placed in a tall plastic cylinder located on a hotplate, with a surface heated to 55 °C (IITC Life Science). Reactions to the heated surface, including hindpaw lick, vocalization or jumping, led to immediate removal from the hotplate. Measures were taken of latency to respond. The maximum test length was 30 s, to avoid paw damage. A two-way ANOVA indicated a significant main effect of sex ($F_{(1,1)} = 8.83$, $P = 0.0053$), and genotype × sex interaction ($F_{(1,35)} = 4.3$, $P = 0.0455$) (Supplementary Table 11). Post-hoc comparisons revealed that the male GPR68-knockout mice had significantly lower latencies to respond than female knockout mice.

Acoustic startle method. The acoustic startle test can be used to assess auditory function and sensorimotor gating. The test is based on the measurement of the reflexive whole-body flinch, or startle response, that follows exposure to a sudden noise. Mice can be evaluated for levels of startle magnitude and prepulse inhibition, which occurs when a weak prestimulus leads to a reduced startle in response to a subsequent louder noise. For this study, animals were tested with a San Diego Instruments SR-Lab system. In brief, mice were placed in a small Plexiglas cylinder within a larger, sound-attenuating chamber. The cylinder was seated upon a piezoelectric transducer, which allowed vibrations to be quantified and displayed on a computer. The chamber included a house light, fan, and a loudspeaker for the acoustic stimuli. Background sound levels (70 dB) and calibration of the acoustic stimuli were confirmed with a digital sound level meter (San Diego Instruments). Each session consisted of 42 trials, which began with a 5-min habituation period. There were seven different types of trials: the no-stimulus trials, trials with the acoustic startle stimulus (40 ms; 120 dB) alone, and trials in which a prepulse stimulus (20 ms; 74, 78, 82, 86 or 90 dB) occurred 100 ms before the onset of the startle stimulus. Measures were taken of the startle amplitude for each trial across a 65-ms sampling window, and an overall analysis was performed for each subject's data for levels of prepulse inhibition at each prepulse sound level (calculated as 100−(response amplitude for prepulse stimulus and startle stimulus together/response amplitude for startle stimulus alone) ×100).

Results from acoustic startle test. The GPR68-knockout mice had decreased startle responses after presentation of acoustic stimuli, in comparison to the wild-type mice (Extended Data Fig. 11e, f). A repeated-measures ANOVA, conducted on startle response amplitudes, indicated significant main effects of genotype ($F(1,35) = 7.22$, $P = 0.011$) and sex ($F(1,35) = 16.61$, $P = 0.0003$), and a genotype × decibel level interaction ($F(6,210) = 5.77$, $P < 0.0001$). Separate comparisons confirmed that both male and female knockout mice showed significant reductions in startle responses (genotype × decibel level interaction, males, $F(6,84) = 2.57$, $P = 0.0245$; and females, $F(6,126) = 3.48$, $P = 0.0032$). The decreased startle responses and overt sex differences were not associated with changes in prepulse inhibition (Extended Data Fig. 11g, h). The significant main effects of genotype on startle were no longer evident during a second acoustic startle test, conducted when mice were 16–17 weeks in age.

Morris water maze, visible platform test. The Morris water maze task was used to assess spatial learning and visual function in the mice. The water maze consisted of a large circular pool (diameter = 122 cm) partially filled with water (45 cm deep, 24–26 °C), located in a room with numerous visual cues. Mice were first tested using a visible platform. In this case, each animal was given four trials per day, across 2 days, to swim to an escape platform cued by a patterned cylinder extending above the surface of the water. For each trial, the mouse was placed in the pool at one of four possible locations (randomly ordered), and then given 60 s to find the visible platform. If the mouse found the platform, the trial ended, and the animal was allowed to remain 10 s on the platform before the next trial began. If the platform was not found, the mouse was placed on the platform for 10 s, and then given the next trial. Measures were taken of latency to find the platform via an automated tracking system (Noldus Ethovision). As shown in Supplementary Table 12, all groups of mice demonstrated a high degree of proficiency in the visual cue task.

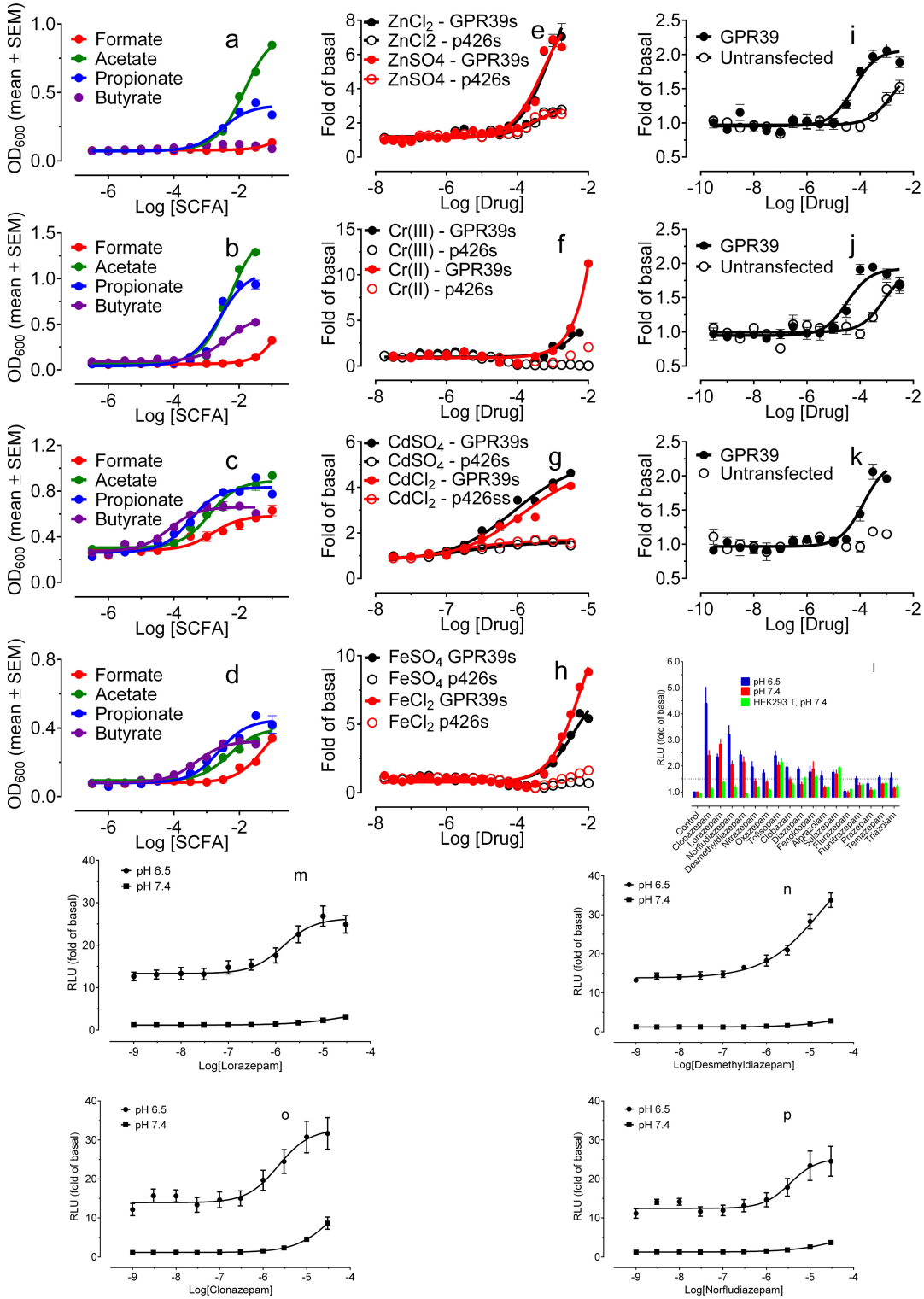
Acquisition and reversal learning in the hidden platform test (Extended Data Fig. 11i–l). Three days after the visual cue task, mice were tested for their ability to find a submerged, hidden escape platform (diameter = 12 cm). As in the procedure for visual cue learning, each animal was given four trials per day, with 1-min per trial, to swim to the hidden platform. The criterion for learning was an average latency of 15 s or less to locate the platform on 1 day. Mice were tested until the criterion was reached, with a maximum of 9 days of testing. When criterion was reached, mice were given a 1-min probe trial in the pool with the platform removed. In this case, selective quadrant search was evaluated by measuring number of crosses over the location where the platform (the target) had been placed during training, and the corresponding areas in the other three quadrants. After the acquisition phase, mice were tested for reversal learning, using the same procedure as described above. In this phase, the hidden platform was located in a different quadrant in the pool, diagonal to its previous location. As before, measures were taken of latency to find the platform. On the day that the criterion for learning was met, the platform was removed from the pool, and the group was given a probe trial to evaluate reversal learning.

For the above behavioural profiling studies, subjects were 21 wild-type mice (9 males and 12 females) and 18 GPR68-knockout mice (7 males and 11 females), on a C57BL/6 background. Sample sizes were not statistically predetermined. Testing began when animals were 6–7 weeks of age. For each procedure, measures were taken by an observer blinded to mouse genotype (wild type or knockout) and no animals were excluded from analysis. Data were analysed using one-way or repeated-measures ANOVA. Fisher's protected least-significant difference tests were used for comparing group means only when a significant F value was determined. Within-group comparisons were conducted to determine side preference in the social behaviour tests. For all comparisons, significance was pre-set at $P < 0.05$.

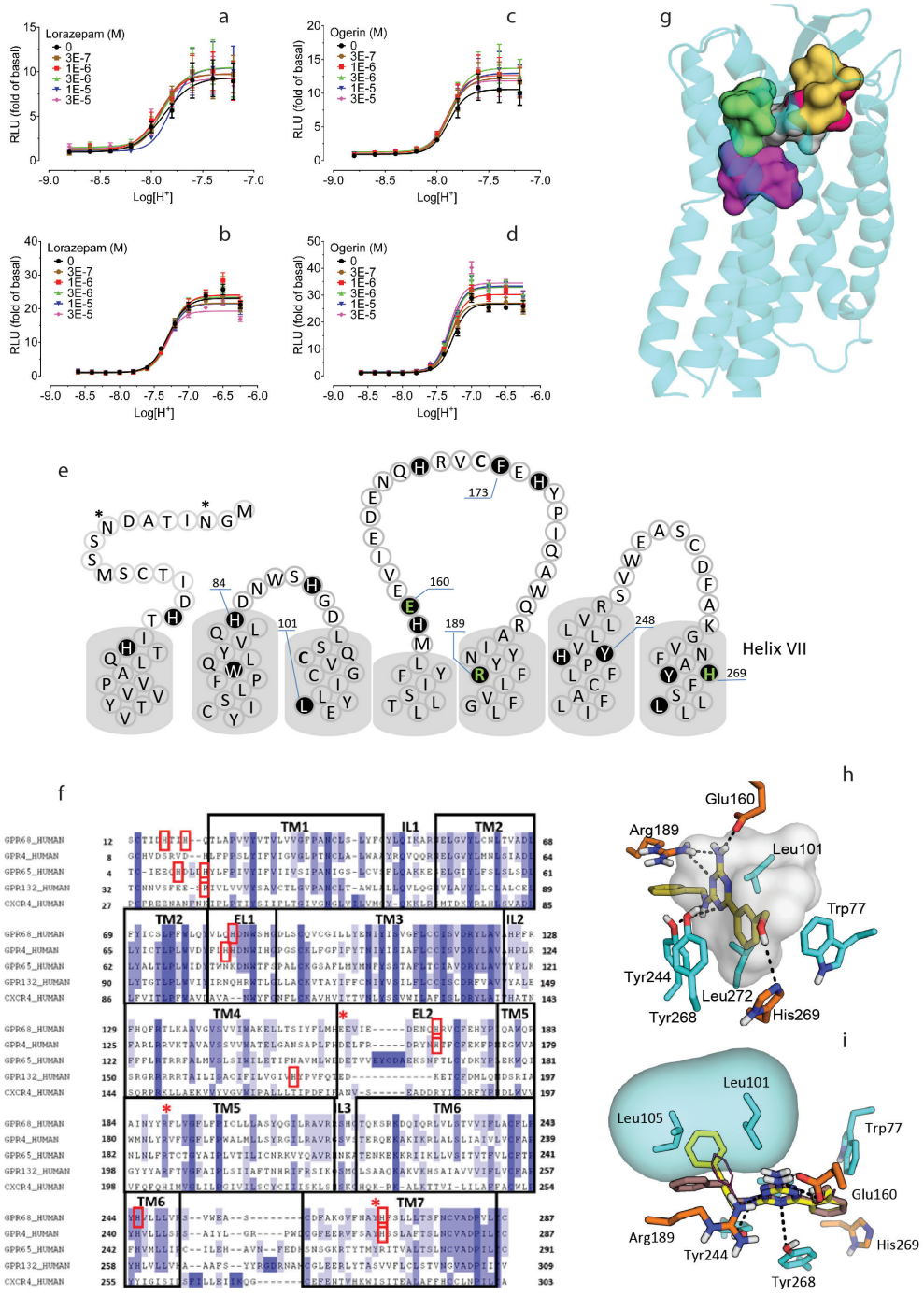
Effect of ogerin and its analogue ZINC32547799 on learning and memory. Contextual and cue-dependent learning and memory were evaluated using a Near-Infrared Video Fear Conditioning system (MED Associates). Test chambers (29 × 25 × 25 cm) had transparent walls and metal rod floors, and were enclosed in sound-attenuating boxes. The conditioned fear procedure had three phases: training, a test for contextual learning, and a test for cue-dependent learning. Before each phase, mice were moved to a holding room adjacent to the test room and acclimated for at least 30 min. In the 8-min training phase, mice receive three pairings of a 30-s, 90-dB, 5-kHz tone (the conditioned stimulus) and a 2-s, 0.6-mA foot shock (the unconditioned stimulus), in which the shock was presented during the last 2 s of the tone. Context-dependent learning was evaluated 24 h after the training phase. Mice were placed back into the original test chamber, and levels of freezing (immobility) were determined across a 5-min session, without the presence of the conditioned or unconditioned stimulus. Forty-eight hours after the training phase, mice were evaluated for associative learning to the auditory cue (the conditioned stimulus) in a final 6-min session. The conditioning chambers were modified using a Plexiglas insert to change the wall and floor surface, and a novel odour (vanilla flavouring) was added to the sound-attenuating box. Baseline behaviour was scored for 2 min, and then three 30-s conditioned stimulus tones were presented across a 4-min period. Levels of freezing were automatically measured by the image tracking software (Med Associates). Freezing was defined as no movement (below the movement threshold) for 0.5 s. To evaluate the effect of drug, strain-matched group of animals were given ogerin (10 mg kg^{-1} in 10% Tween 80 or saline) 30 min before the training.

For the learning and memory studies, sample sizes (number of animals) were not predetermined by a statistical method, and minimum of six male animals (age of 6–8 weeks) were used in each group (exact number of animals was specified in figure legends). Animals were assigned to groups randomly and experiments were not blinded to investigators. No animals were excluded from analysis. Statistical analyses were performed after first assessing the normality of distributions of data sets. Comparisons between groups were made using unpaired t-tests. Welch's corrections were used when variances between groups were unequal. Comparisons between groups during conditioning, contextual and cued memory tests were assessed using two-way ANOVA with $P < 0.05$ being considered significant.

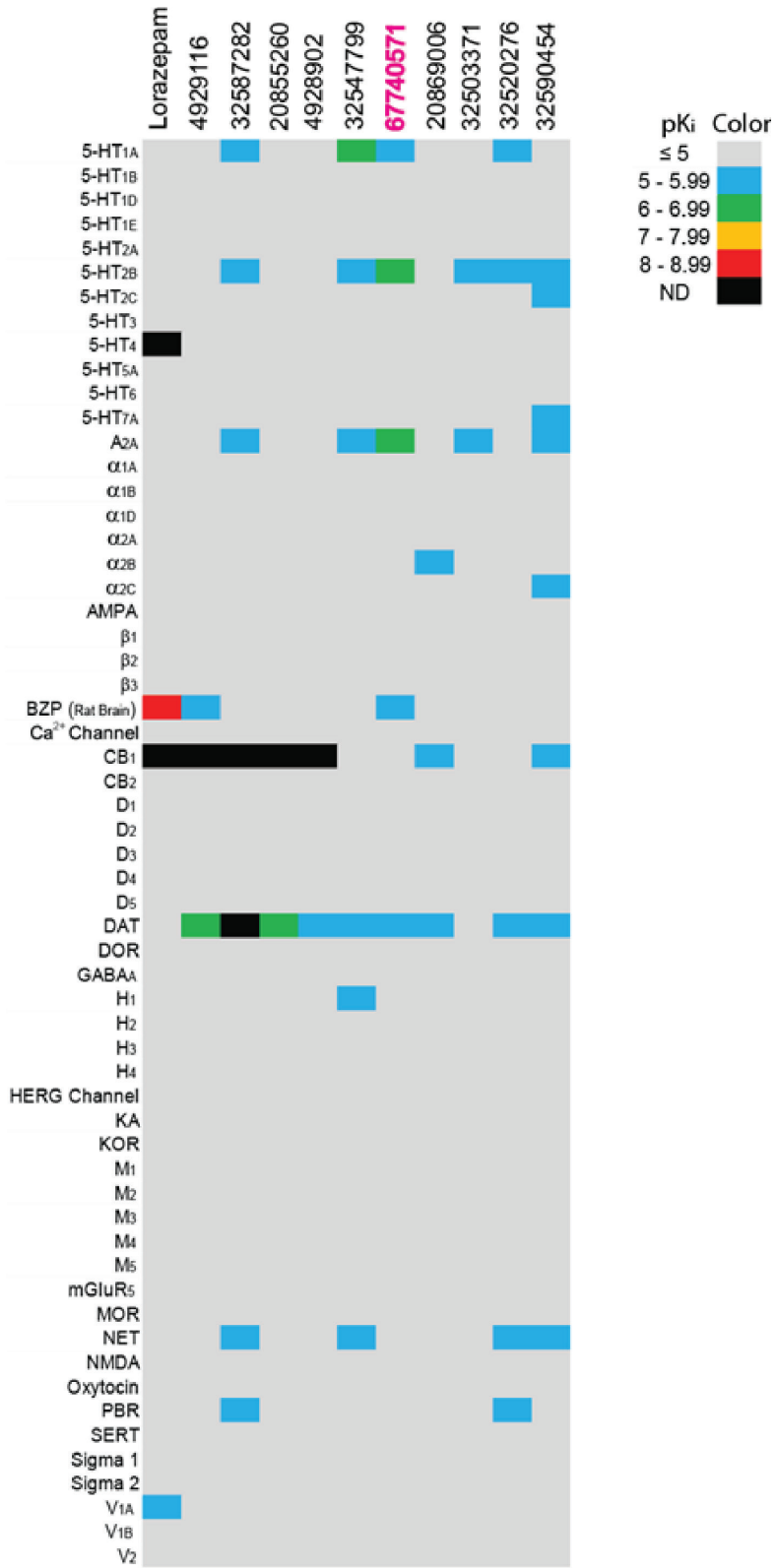
45. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **Chapter 2,** Unit 2.9 (2001).
46. Yang, Q. & Sharp, K. A. Building alternate protein structures using the elastic network model. *Proteins* **74,** 682–700 (2009).
47. Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **320,** 597–608 (2002).
48. Li, J., Zhu, T., Cramer, C. J. & Truhlar, D. G. New class IV charge model for extracting accurate partial charges from wave functions. *J. Phys. Chem. A* **102,** 1820–1831 (1998).
49. Chambers, C. C., Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *J. Phys. Chem.* **100,** 16385–16398 (1996).
50. Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I. & Shoichet, B. K. Quantifying the relationships among drug classes. *J. Chem. Inf. Model.* **48,** 755–765 (2008).
51. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40,** D1100–D1107 (2012).
52. Mumberg, D., Muller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156,** 119–122 (1995).
53. Erlenbach, I. *et al.* Functional expression of M_1, M_3 and M_5 muscarinic acetylcholine receptors in yeast. *J. Neurochem.* **77,** 1327–1337 (2001).
54. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl Acad. Sci. USA* **104,** 5163–5168 (2007).
55. Christopoulos, A. & Kenakin, T. G protein-coupled receptor allosterism and complexing. *Pharmacol. Rev.* **54,** 323–374 (2002).
56. Besnard, J. *et al.* Automated design of ligands to polypharmacological profiles. *Nature* **492,** 215–220 (2012).
57. Keiser, M. J. *et al.* Predicting new molecular targets for known drugs. *Nature* **462,** 175–181 (2009).
58. Horvat, S. J. *et al.* A-kinase anchoring proteins regulate compartmentalized cAMP signaling in airway smooth muscle. *FASEB J.* **26,** 3670–3679 (2012).
59. Huang, X.-P., Mangano, T., Hufeisen, S., Setola, V. & Roth, B. L. Identification of human *ether-à-go-go* related gene modulators by three screening platforms in an academic drug-discovery setting. *Assay Drug Dev. Technol.* **8,** 727–742 (2010).



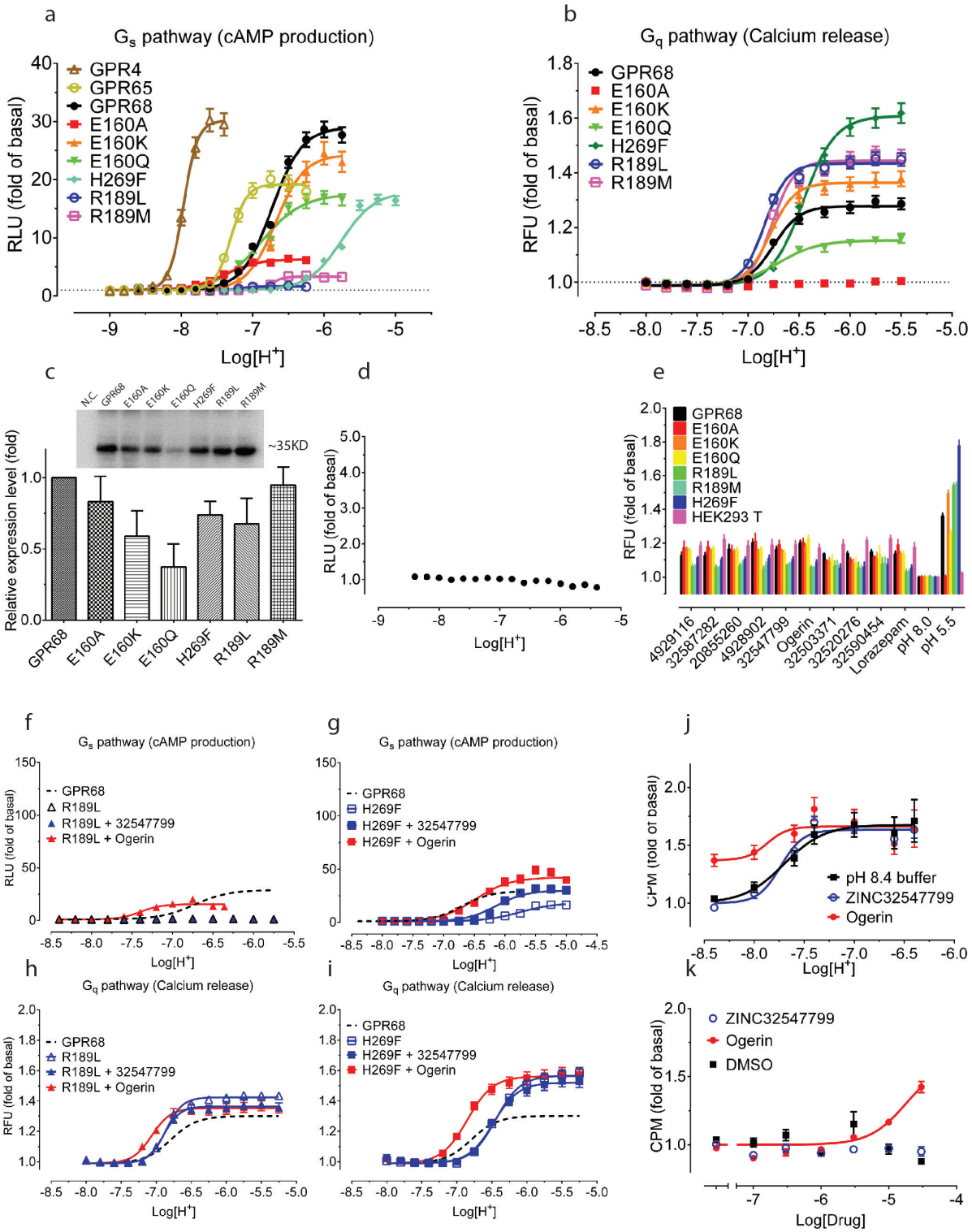
Extended Data Figure 1 | Validation and confirmation of GPCR activation assays. a–o, Yeast (**a–k**) and HEK293T cell (**l–o**) GPCR activation assays. **a–d**, Concentration-dependent growth of GPR43-expressing G_s yeast (**a**), GPR43-expressing G_q yeast (**b**), GPR41-expressing G_s yeast (**c**), and GPR41-expressing G_q yeast (**d**) in response to various short-chain fatty acids (SCFAs). **e–h**, Concentration-dependent growth of GPR39-expressing G_s yeast ($GPR39_s$) in response to zinc ions (**e**), chromium ions (**f**), cadmium ions (**g**) and iron ions (**h**). **i–k**, Concentration-dependent cAMP responses of GPR39-expressing HEK293T cells to $ZnCl_2$ (**i**), $ZnSO_4$ (**j**) or $CdSO_4$ (**k**) as measured by luciferase cAMP reporter assay. **l**, *N*-unsubstituted benzodiazepines (lorazepam, clonazepam, desmethyldiazepam and norfludiazepam; 10 μM) stimulated cAMP production in a GPR68- and pH-dependent manner. Data are mean ± s.e.m. ($n = 3$–66 measurements). **m–p**, Concentration–response curves of *N*-unsubstituted benzodiazepines lorazepam (**m**), desmethyldiazepam (**n**), clonazepam (**o**) and norfludiazepam (**p**) at pH 6.50 or 7.40 in GPR68-transfected HEK293T cells (structures in Supplementary Table 1). Normalized results represent mean ± s.e.m. ($n = 3$) and curves were analysed in GraphPad Prism using the built-in 4 parameter logistic function.



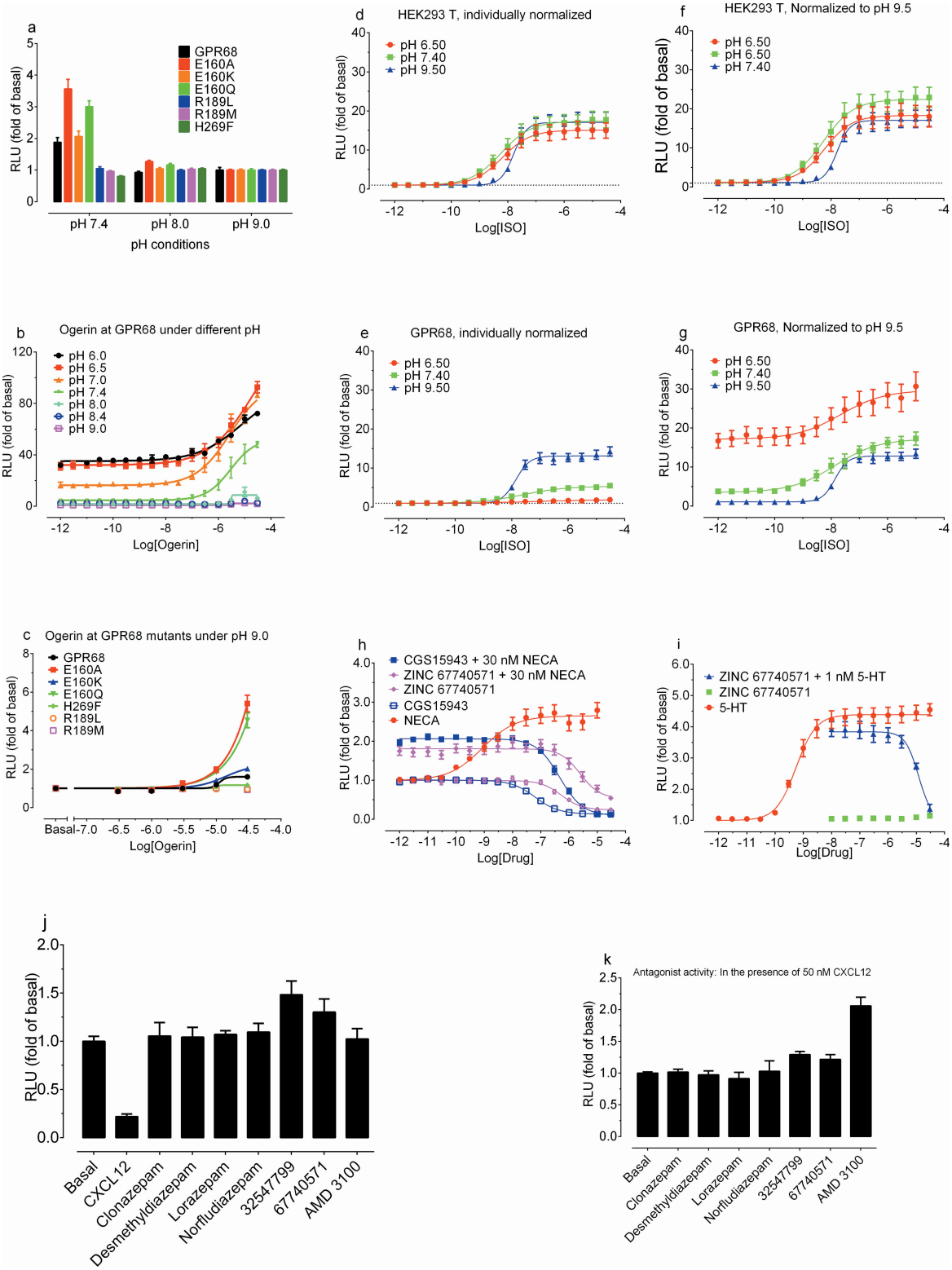
Extended Data Figure 2 | Lorazepam and ogerin have minimal GPR4 or GPR65 activity. a–d, Effect of lorazepam (**a**, **b**) or ogerin (**c**, **d**) on GPR4 (**a**, **c**) or GPR65 (**b**, **d**); data represent normalized mean ± s.e.m. ($n = 3$). Sequencing alignment proton-sensing receptor and docking poses for ogerin and its analogues. **e**, GPR68 snake plot showing extracellular loops and transmembrane domains (upper portion); important residues are highlighted. Glu160, Arg189 and His269 were mutated in this study. **f**, Sequence alignment of GPR4, GPR65 and GPR68 to CXCR4 (PDB code 3ODU) (PROMALS-3D) was manually refined to reduce gaps and to position conserved residues. TM, transmembrane regions; IL, intracellular loop; EL, extracellular loop. Conserved residues highlighted in blue by degree of conservation while red boxes indicate residues important for receptor function. Red stars indicate residues mutated in this study. **g**, Sampling different regions for lorazepam binding modes in GPR68. Yellow and grey surfaces contour the binding site of 1T1t and CVX15 in CXCR4 crystal structures (PDB codes 3ODU and 3OE0, respectively), while green and red surfaces sample the entire binding pocket. The magenta surface represents the canonical orthosteric biogenic amine site. **h**, ZINC32547799 in its predicted orientation and interactions with GPR68. **i**, Optimization of ogerin (magenta, thin lines) to **C2** (brown, structure in Fig. 3a) by insertion of a single methylene is predicted to improve packing in the aryl pocket of the ogerin site. Adding a second methylene, thus creating a propyl linker in **C3** (yellow, structure in Fig. 3a), is predicted to disrupt the packing and thus to reduce the allosteric effect.



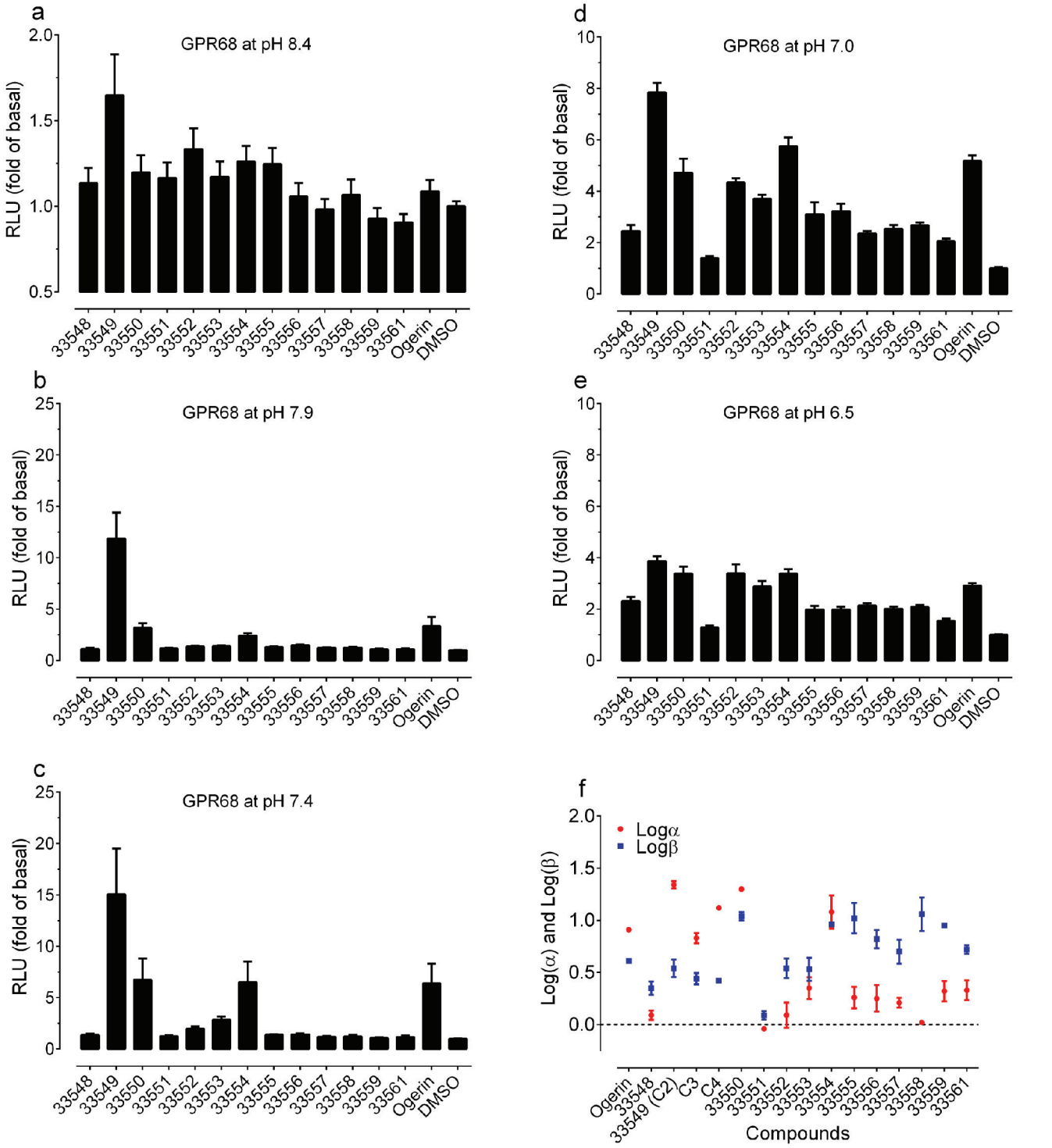
Extended Data Figure 3 | Heat map of off-target activities of lead compounds at potential CNS drug targets. Radioligand binding assays were carried out by the National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP) as described previously[56,57] (online protocols available at http://pdsp.med.unc.edu/pdspw/binding.php). Values represent mean binding affinities (pK_i, $n = 2–4$). Affinities lower than a pK_i of 5, or less than 50% inhibition at 10 μM, are shown as a minimum of 5 on the pK_i scale. The hERG inhibition activity was tested in a hERG functional assay as previously published[59]. AMPA, aminomethylphosphonic acid receptor; BZP, benzodiazepine receptor; DAT, dopamine transporter; DOR, delta (δ) opioid receptor; KA, kainate acid receptor; KOR, kappa (κ) opioid receptor; MOR, mu (μ) opioid receptor; NAT, noradrenaline transporter; NMDA, *N*-methyl-D-aspartate receptor; ND, not determined; PBR, peripheral benzodiazepine binding site; SERT, serotonin transporter; hERG, human ether-a-go-go-related gene (potassium channel Kv11.1).



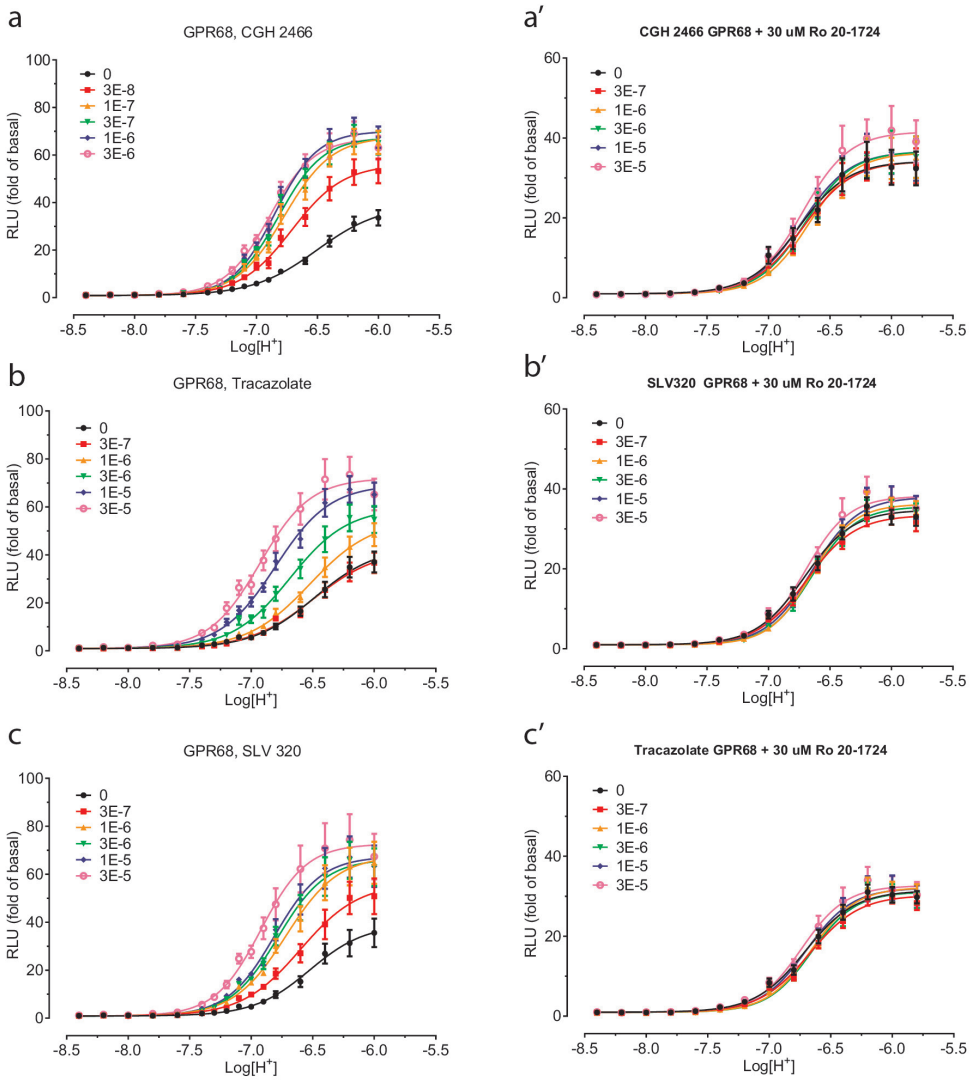
Extended Data Figure 4 | Confirmation of modelling results via mutagenesis. a, **b**, Protons showed agonist activity at GPR68 wild-type and mutant receptors in cAMP production (**a**) and calcium release (**b**); parameters are in Supplementary Table 4. **c**, Relative GPR68 wild-type and mutant receptor expression levels determined by anti-Flag immunoblotting ($n = 3$). **d**, Proton-mediated cAMP production in untransfected cells ($n = 16$). **e**, Calcium release by lorazepam and selected ZINC compounds (10 μM at pH 8.0, $n = 6$–22 measurements). **f**–**j**, Effect of ogerin and ZINC32547799 (10 μM) on proton-mediated cAMP production (**f** and **g**, $n = 4$), calcium release (**h** and **i**, $n = 3$), and phosphatidylinositol hydrolysis (**j**, $n = 3$) at GPR68 wild-type or mutant-transfected HEK293T cells. **k**, Effect of ogerin and ZINC32547799 on phosphatidylinositol hydrolysis at pH 8.4 at GPR68-transfected GPR68 HEK293T cells ($n = 3$). Normalized results represent mean ± s.e.m. and curves were analysed using a four-parameter logistic function.

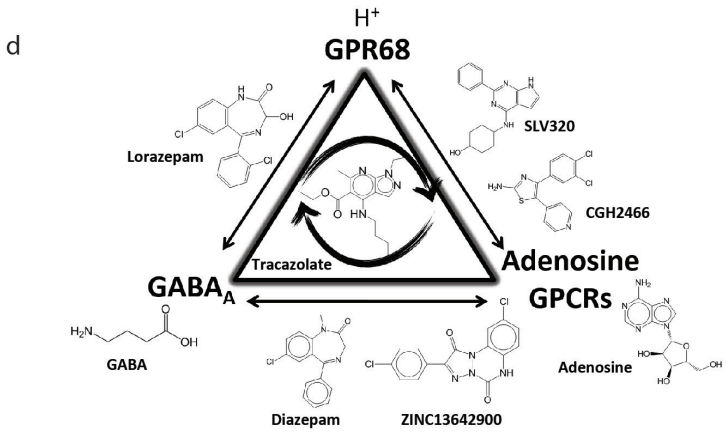


Extended Data Figure 5 | Control experiments for signalling and pharmacology. a, Basal cAMP production of GPR68 wild-type and mutant receptors (mean ± s.e.m., $n = 24–46$ measurements). **b**, pH-dependent activity of ogerin at GPR68 wild type (mean ± s.e.m., $n = 3$). **c**, Ogerin concentration–responses at GPR68 wild-type and mutant receptors at pH 9.0 (**c**, mean ± s.e.m., $n = 3$), under which cAMP reporter assay was *n*ot affected (**d**–**f**). **d**–**g**, Proton modulated isoproterenol-mediated G_s-activation via β_2-adrenergic receptors in untransfected (**d**, **f**) and GPR68-transfected (**e**, **g**) cells. Normalized results (basal at pH 9.5 for **d** and **e**; or corresponding buffer control for **f** and **g**) represent mean ± s.e.m. ($n = 6$). **h**, **i**, Inverse agonist and antagonist activity (K_i of 220 nM) of ogerin at A_{2A} (cAMP production, **h**) and weak antagonist activity (K_i of 736 nM) at 5-HT_{2B} receptors (calcium mobilization, **i**). 5′-*N*-ethylcarboxamidoadenosine (NECA) and 2-chloro-N^6-cyclopentyladenosine (CCPA) served as agonist controls, while CGS15943 is an inverse agonist control for A_{2A} receptors. Normalized results represent mean ± s.e.m. ($n = 3$). Curves were analysed in GraphPad Prism with the built-in four-parameter logistic function. **j**, **k**, Lead compounds (10 μM) showed no agonist (**j**) or antagonist (**k**) activity at CXCR4 receptors (cAMP production) with CXCL12 as an agonist control (1 or 3 μM) or AMD 3100 (10 μM) as an antagonist control. Results represent mean ± s.d. ($n = 2$).

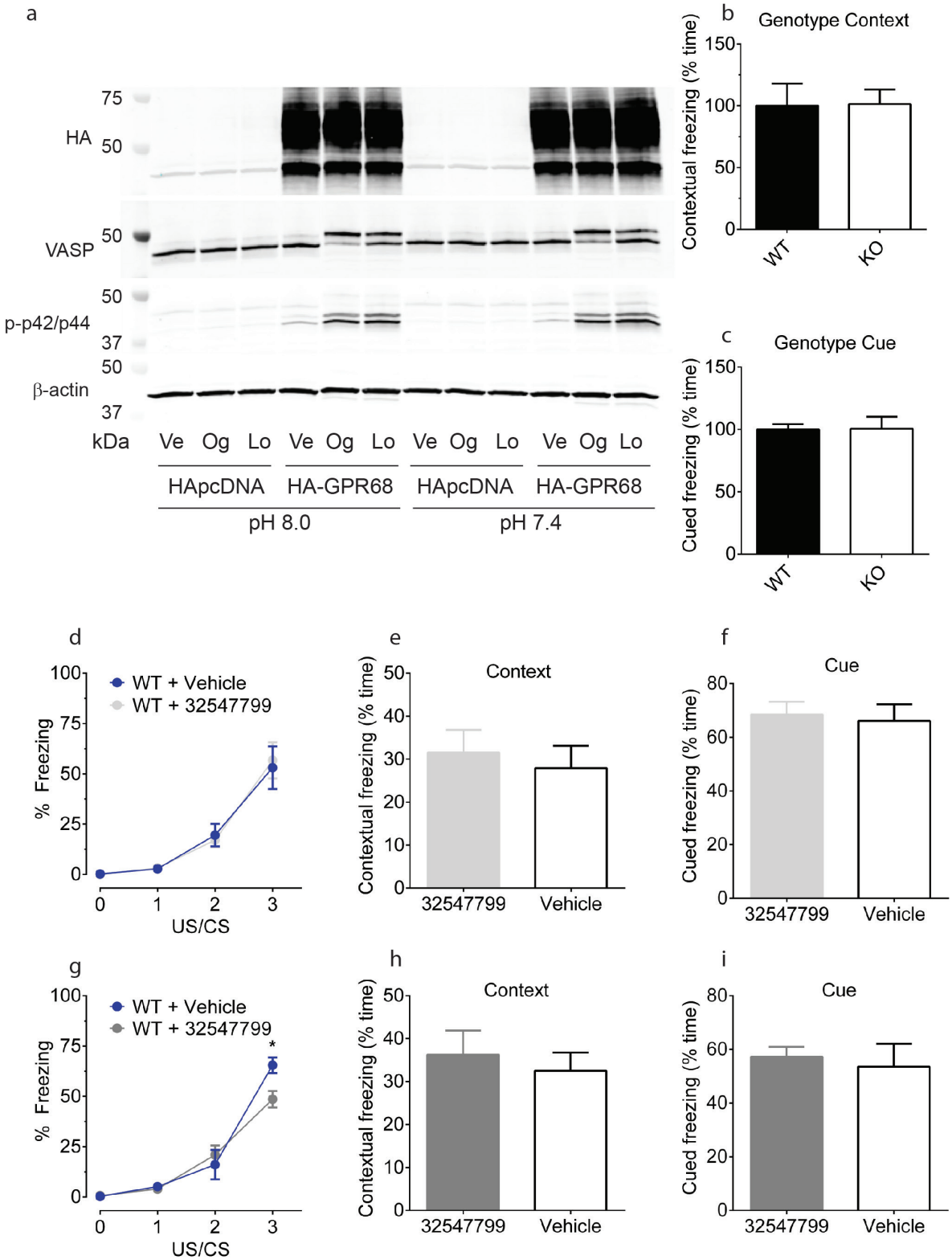


Extended Data Figure 6 | Primary screening and comparison of allosteric parameters of 13 ogerin analogues at GPR68. The 13 ogerin analogues (structures in Supplementary Table 9) identified from docking a virtual library of more than 600 ogerin derivatives were synthesized (Supplementary Information). **a**–**e**, Production of cAMP was measured in transiently transfected HEK293T cells at 10 μM and five different pH conditions, pH 8.4 (**a**); pH 7.9 (**b**); pH 7.4 (**c**); pH 7.0 (**d**); and pH 6.5 (**e**), to reveal any pH-dependent potentiation activity. Normalized results represent mean ± s.e.m. ($n = 8$–16 measurements). **f**, Graphic comparison of the allosteric parameters logα and logβ. Proton concentration–responses were carried out in the absence and presence of increasing concentrations of ogerin and its analogues, results were analysed using a standard allosteric operational model to obtain allosteric parameters. Values represent mean ± s.e.m. ($n \geq 3$; see details in Supplementary Table 8).

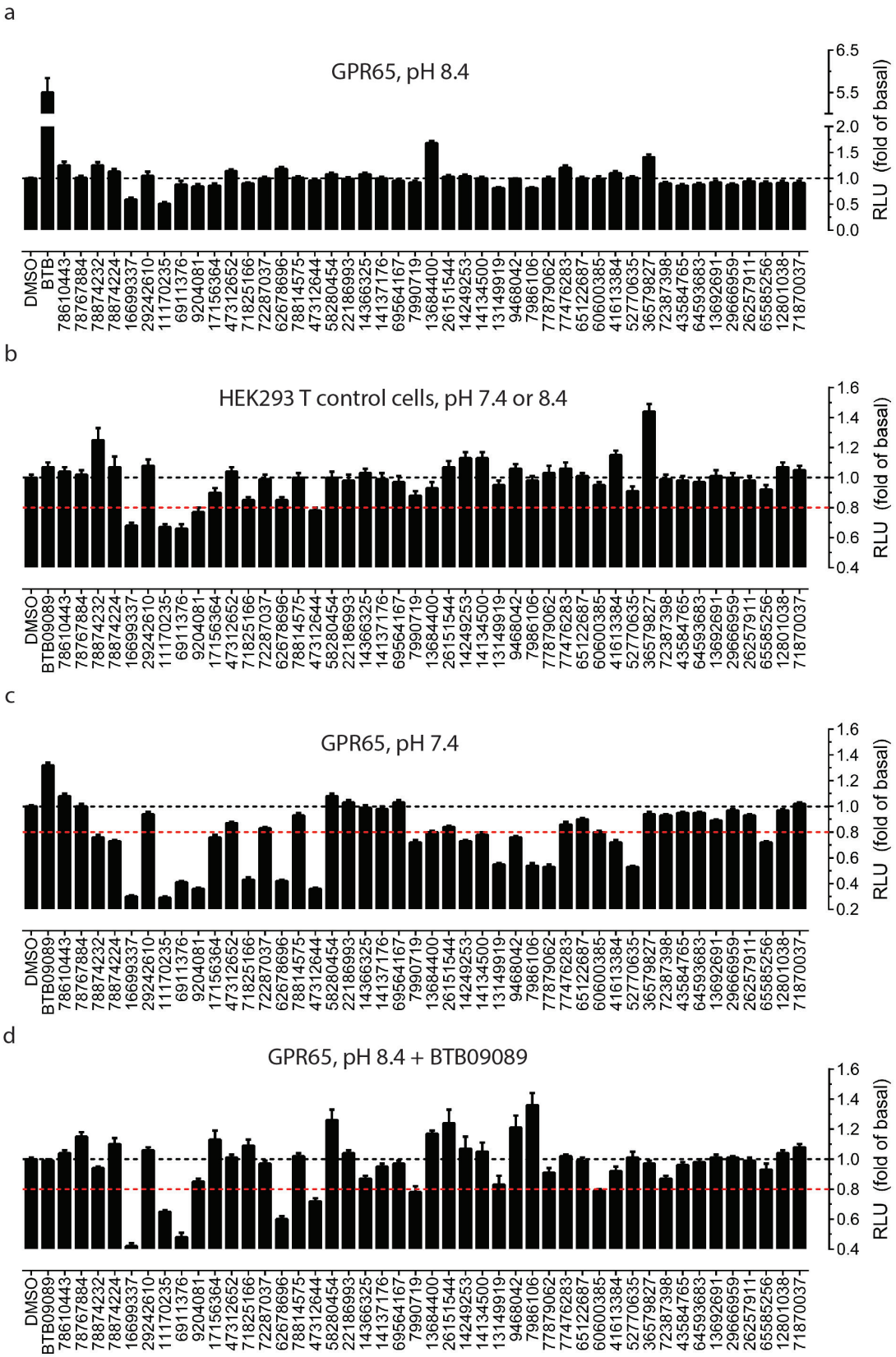




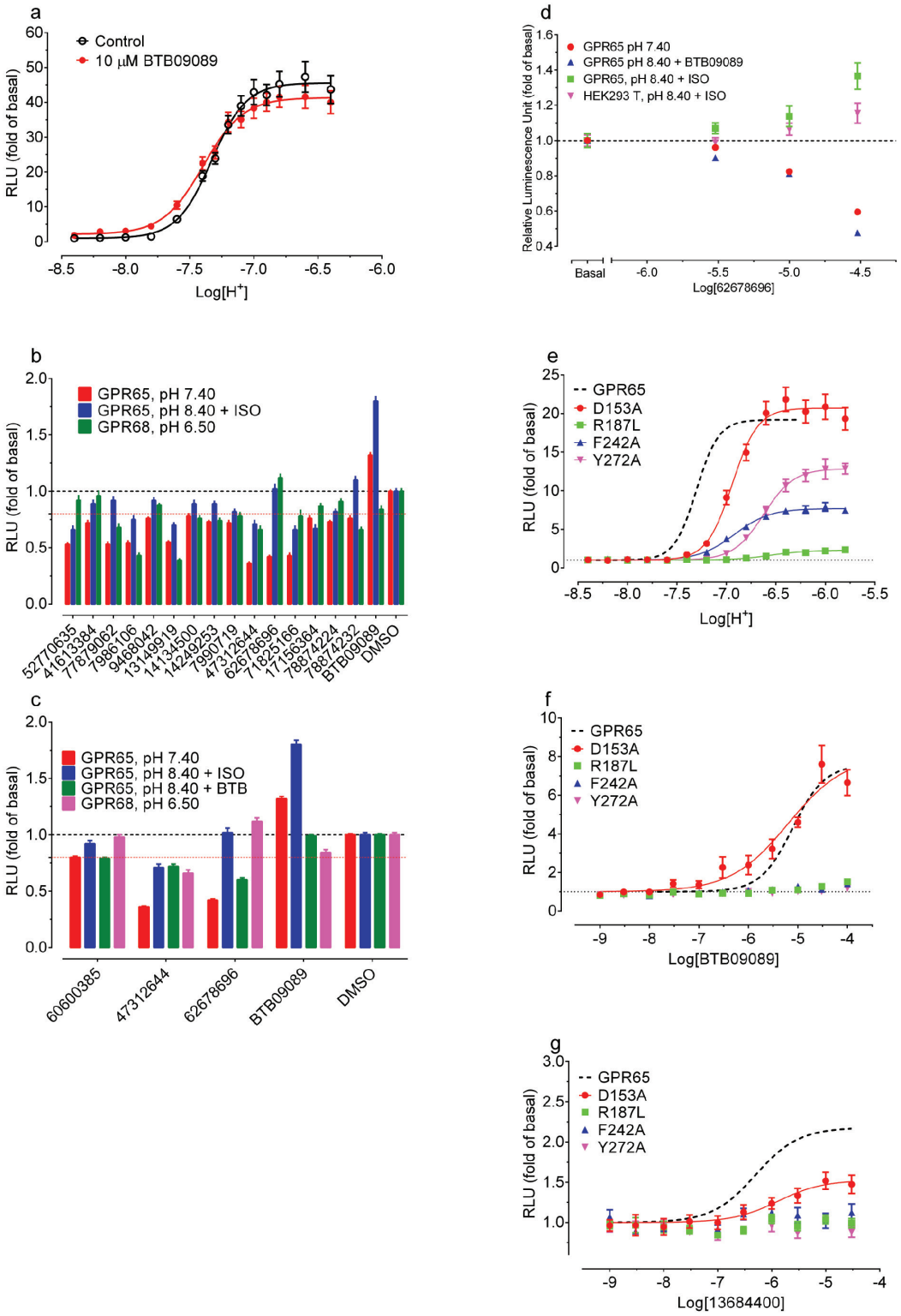
Extended Data Figure 7 | Characterization of potent GPR68 PAMs. **a**–**c**, Concentration–response curves of H^+ in the absence and presence of increasing concentrations of CGH2466 (**a**, **a′**), tracazolate (**b**, **b′**) and SLV320 (**c**, **c′**) and in the absence (left column, **a**, **b**, **c**) and presence (right column, **a′**, **b′**, **c′**) of phosphodiesterase inhibitor (Ro 20-1724, 30 μM) at GPR68-expressing cells. Normalized results (mean ± s.e.m., $n = 8$ for CGH2466; $n = 5$ for tracazolate; $n = 5$ for SLV320 for left column and $n = 3$ for right column) were analysed using a four-parameter logistic function and the standard allosteric operational model (not shown). Allosteric parameters in absence of Ro 20-1724 are summarized in Supplementary Table 8. For each pair of fittings, the proton potency value (negative logarithm of the half-maximum effective concentration (pEC_{50})) from the agonist concentration–response curve (right) in the absence of testing compound was used as the pK_A for the allosteric operational model (left). **d**, Schematic showing the shared pharmacology among $GABA_A$, adenosine GPCRs and GPR68 ligands. Molecules along each edge of the triangle have been shown to have activity at both targets, whereas tracazolate, in the middle, shows activity at all three.



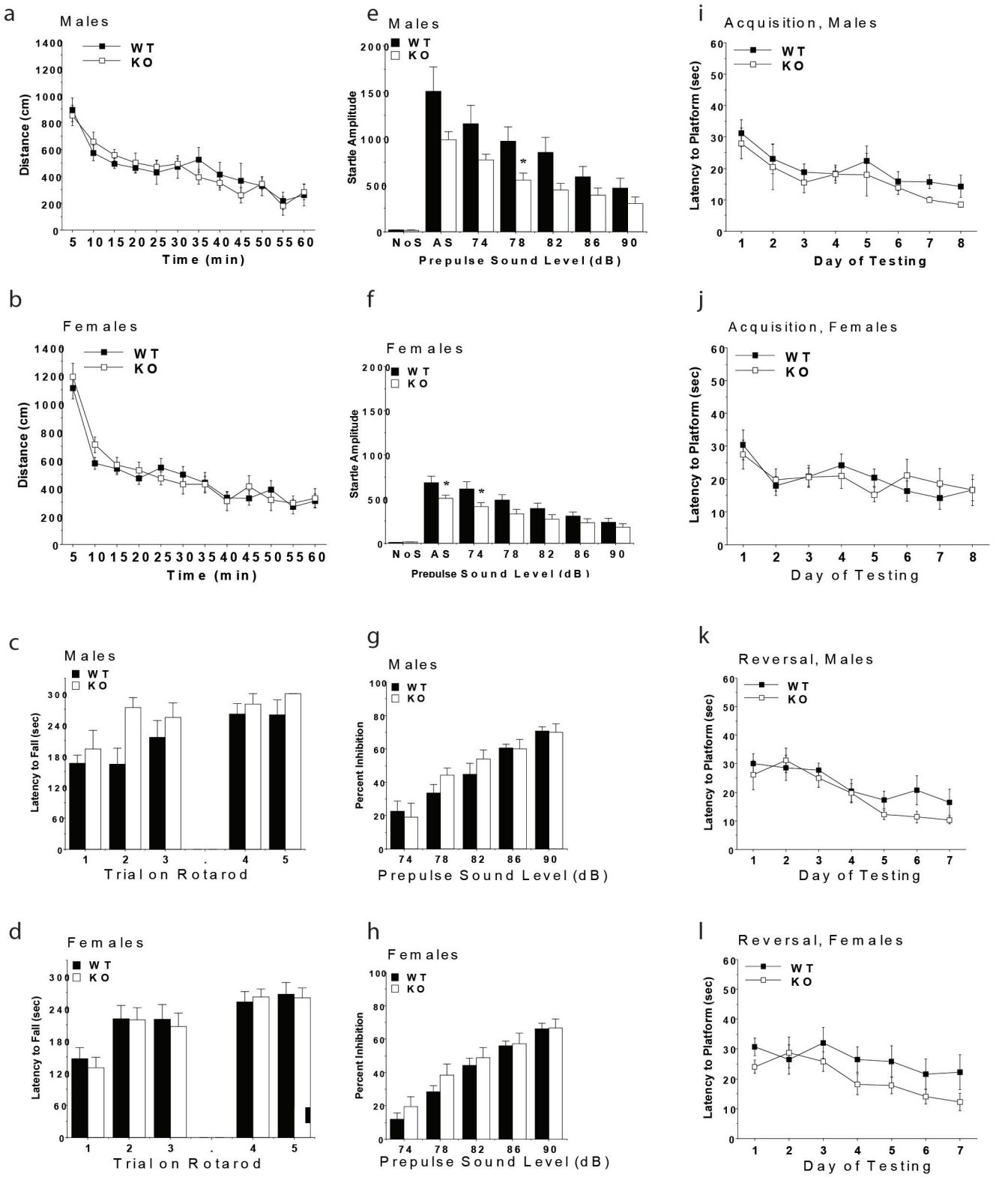
Extended Data Figure 8 | GPR68 mouse biology. a, Ogerin (Og) and lorazepam (Lo) activate PKA and p42/p44 MAP kinase in HEK293 cells stably expressing haemagglutinin (HA)-tagged GPR68 but not HApcDNA; vehicle (Ve). **b**, **c**, GPR68 knockout ($n = 7$) mice exhibited no differences in contextual memory retrieval (**b**) or cued memory retrieval (**c**) as compared to wild-type mice ($n = 8$). **d**–**f**, At 10 mg kg^{-1}, the ogerin isomer ZINC32547799 had no effect on learning (**d**) or contextual and cue memory (**e**, **f**) in wild-type mice (vehicle, $n = 6$; drug, $n = 7$). **g**–**i**, At 30 mg kg^{-1}, ZINC32547799 enhanced wild-type learning (**g**, drug × time interaction, $F_{(3.39)} = 3.58$, $P = 0.022$; drug alone $F_{(1.39)} = 1.19$, $P = 0.295$; Bonferroni post-hoc test revealed a significant effect ($P < 0.05$) at the third unconditioned/conditioned stimulus training point, two-way ANOVA), but had no effect at contextual and cue memory (**h**, **i**) (vehicle, $n = 7$; drug $n = 8$). Male mice at age of 6–8 weeks were used in the test. Normalized contextual memory retrieval (**d**) and cued memory retrieval (**f**) are presented in Fig. 4c, d.



Extended Data Figure 9 | Screening of ZINC compounds predicted to be active at GPR65 based on BTB09089 docking poses. a–d, Primary screening with ZINC compounds (30 μM) for agonist activity at GPR65 when receptors were kept inactive at pH 8.40 (**a**); at control HEK293T cells for nonspecific activity (**b**); at GPR65 when receptors were activated at pH7.40 for modulator or antagonist activity (**c**); at GPR65 when receptors were activated by BTB09089 (30 μM) at pH 8.40 for modulator or antagonist activity (**d**). Normalized results represent mean ± s.e.m. from a minimum of three assays (each in minimum of triplicate and a total of ≥16 measurements). The red dashed line in **b–d** indicates the 20% inhibition line (an arbitrary cut-off line).



Extended Data Figure 10 | Characterization of GPR65 allosteric modulators at wild-type and mutant receptors. a, BTB09089 showed weak agonist activity, but failed to potentiate proton activity at GPR65 ($n = 8$). **b**, Selected compounds from Extended Data Fig. 9b, c were tested for GPR65 specific inhibition ($n = 16$–56 measurements). Several compounds (such as ZINC41613384, ZINC9468042 and ZINC62678696) showed GPR65-specific inhibition. **c**, Selected compounds from Extended Data Fig. 9b, d were tested for antagonist activity against BTB09089-activated signal at GPR65 ($n = 16$–64 measurements). ZINC62678696 showed GPR65 specific inhibition when it was activated by either proton or BTB09089. **d**, ZINC62678696 inhibited GPR65 activity. **e**–**g**, Proton concentration–responses (**e**), BTB09089 concentration–responses (**f**), and ZINC13684400 concentration–responses (**g**) at GPR65 mutant receptors. Normalized results represent mean ± s.e.m. ($n \geq 3$) and curves were analysed in GraphPad Prism with a standard four-parameter logistic function. Corresponding curves of proton at GPR65 wild-type receptors (from Extended Data Fig. 4a) and BTB09089 and ZINC13684400 (from Fig. 5e) are also included (dashed lines) for comparison. Pharmacological parameters are listed in Supplementary Table 13.



Extended Data Figure 11 | *In vivo* behavioural profiling of GPR68-knockout mice. a, **b**, No effects of GPR68 deletion on distance travelled in an open field. Data represent mean ± s.e.m. for each group for a one-hour test session. **c**, **d**, No difference on latency to fall from an accelerating rotarod. Data represent mean ± s.e.m. for each group. **e**–**h**, Decreased startle responses in GPR68 knockout mice after presentation of acoustic stimuli (**e**, **f**). Data represent mean ± s.e.m. for each group. No effects of genotype were found for levels of prepulse inhibition (**g**, **h**). Data represent mean ± s.e.m. for each group (*$P < 0.05$). **i**–**l**, No difference at acquisition and reversal learning in the Morris water maze. Data represent mean ± s.e.m. of four trials per day. Subject numbers were 9 wild-type and 7 knockout male mice, and 12 wild-type and 11 knockout female mice.