

# Structures of the $\sigma_2$ receptor enable docking for bioactive ligand discovery

<https://doi.org/10.1038/s41586-021-04175-x>

Received: 20 April 2021

Accepted: 19 October 2021

Published online: 08 December 2021

 Check for updates

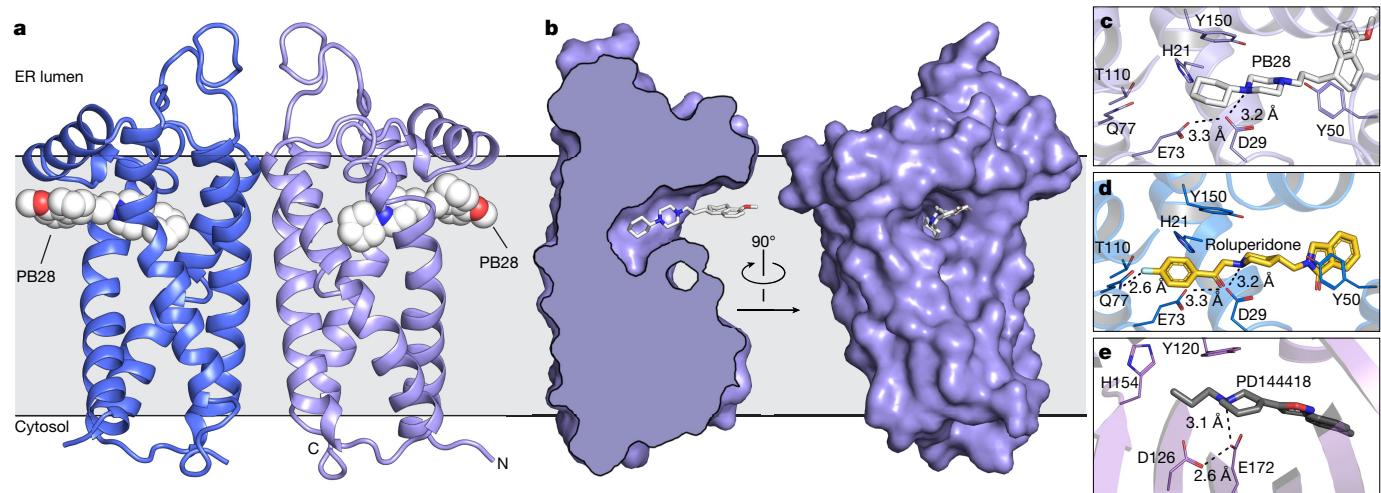
Assaf Alon<sup>1,12</sup>, Jiankun Lyu<sup>2,12</sup>, Joao M. Braz<sup>3,12</sup>, Tia A. Tummino<sup>2,4</sup>, Veronica Craik<sup>3</sup>, Matthew J. O'Meara<sup>5</sup>, Chase M. Webb<sup>2,4</sup>, Dmytro S. Radchenko<sup>6,7</sup>, Yurii S. Moroz<sup>8</sup>, Xi-Ping Huang<sup>9,10</sup>, Yongfeng Liu<sup>9,10</sup>, Bryan L. Roth<sup>9,10,11</sup>, John J. Irwin<sup>2</sup>, Allan I. Basbaum<sup>3,12</sup>, Brian K. Shoichet<sup>2,12</sup> & Andrew C. Kruse<sup>1,12</sup>

The  $\sigma_2$  receptor has attracted intense interest in cancer imaging<sup>1</sup>, psychiatric disease<sup>2</sup>, neuropathic pain<sup>3–5</sup> and other areas of biology<sup>6,7</sup>. Here we determined the crystal structure of this receptor in complex with the clinical candidate roluperidone<sup>2</sup> and the tool compound PB28<sup>8</sup>. These structures templated a large-scale docking screen of 490 million virtual molecules, of which 484 compounds were synthesized and tested. We identified 127 new chemotypes with affinities superior to 1  $\mu$ M, 31 of which had affinities superior to 50 nM. The hit rate fell smoothly and monotonically with docking score. We optimized three hits for potency and selectivity, and achieved affinities that ranged from 3 to 48 nM, with up to 250-fold selectivity versus the  $\sigma_1$  receptor. Crystal structures of two ligands bound to the  $\sigma_2$  receptor confirmed the docked poses. To investigate the contribution of the  $\sigma_2$  receptor in pain, two potent  $\sigma_2$ -selective ligands and one potent  $\sigma_1/\sigma_2$  non-selective ligand were tested for efficacy in a mouse model of neuropathic pain. All three ligands showed time-dependent decreases in mechanical hypersensitivity in the spared nerve injury model<sup>9</sup>, suggesting that the  $\sigma_2$  receptor has a role in nociception. This study illustrates the opportunities for rapid discovery of in vivo probes through structure-based screens of ultra large libraries, enabling study of underexplored areas of biology.

The  $\sigma$  receptors are integral membrane proteins that are widely expressed in both the central nervous system and in peripheral tissues<sup>10</sup>. They are divided into  $\sigma_1$  and  $\sigma_2$  subtypes on the basis of differences in tissue distribution and in pharmacological profile<sup>11</sup>, but despite their names, the two proteins are unrelated in sequence. Cloned in 1996, the  $\sigma_1$  receptor has no parologue within the human genome; its closest homologue of known function is the yeast  $\Delta 8,7$  sterol isomerase ERG2<sup>12</sup>. Studies conducted on  $\sigma_1$ -knockout mouse tissue<sup>13</sup> showed that the  $\sigma_2$  receptor is not a splice variant or modified form of  $\sigma_1$ , but rather derives from an unrelated gene. The molecular identity of the  $\sigma_2$  receptor remained unknown until we purified it from calf liver tissue<sup>14</sup> and showed that it is TMEM97, an endoplasmic reticulum (ER)-resident membrane protein that regulates the sterol transporter NPC1<sup>15,16</sup>. TMEM97 is predicted to be a four-helix bundle protein with both amino and carboxy termini facing the cytoplasm. A member of the EXPERA family<sup>17</sup>, the  $\sigma_2$  receptor is distantly related to emopamil-binding protein (EBP), the mammalian  $\Delta 8,7$  sterol isomerase that is required for cholesterol synthesis, and to TM6SF2, which regulates liver lipid homeostasis<sup>18</sup>.

The  $\sigma_2$  receptor is overexpressed in proliferating cells and in many tumours<sup>19</sup>, and labelled  $\sigma_2$  ligands have been proposed as tools for cancer diagnosis and therapy<sup>1</sup>. A ternary complex between the  $\sigma_2$  receptor, PGRMC1 and the low density lipoprotein (LDL) receptor was reported to increase the rate of LDL internalization<sup>7</sup>. Consistent with its high expression in the central nervous system (CNS), the  $\sigma_2$  receptor has also been proposed as a target for the treatment of CNS disorders. The  $\sigma_2$  receptor ligand Elayta (CT1812) is in clinical trials for mild to moderate Alzheimer's disease<sup>6</sup>, and roluperidone (MIN-101) is in clinical development for schizophrenia<sup>2</sup>. When tested in animal models,  $\sigma_2$  receptor ligands reduce alcohol-withdrawal symptoms<sup>5,20</sup> and have a neuroprotective effect in brain injury<sup>21</sup>. Finally, recent studies have found that  $\sigma_2$  ligands are anti-allodynic in models of neuropathic pain<sup>3–5</sup>. As this is also true of  $\sigma_1$  ligands, and because most  $\sigma_2$  ligands cross-react with the  $\sigma_1$  receptor, probe ligands that are selective for  $\sigma_2$  over  $\sigma_1$  would help to elucidate  $\sigma_2$  biology and could be used in the development of therapeutic agents. However, little is known of the molecular architecture of the  $\sigma_2$  receptor or the structural bases for ligand recognition, which has hindered the discovery of selective ligands<sup>22,23</sup>. Here we used

<sup>1</sup>Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA. <sup>2</sup>Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA. <sup>3</sup>Department of Anatomy, University of California, San Francisco, San Francisco, CA, USA. <sup>4</sup>Graduate Program in Pharmaceutical Sciences and Pharmacogenomics, University of California, San Francisco, San Francisco, CA, USA. <sup>5</sup>Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. <sup>6</sup>Enamine, Kyiv, Ukraine. <sup>7</sup>Taras Shevchenko National University of Kyiv, Kyiv, Ukraine. <sup>8</sup>Chemspace, Kyiv, Ukraine. <sup>9</sup>Department of Pharmacology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA. <sup>10</sup>National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA. <sup>11</sup>Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>12</sup>These authors contributed equally: Assaf Alon, Jiankun Lyu, Joao M. Braz. <sup>13</sup>e-mail: Allan.Basbaum@ucsf.edu; shoichet@cgl.ucsf.edu; Andrew\_Kruse@hms.harvard.edu



**Fig. 1 | Structure of the  $\sigma_2$  receptor and binding site ligand recognition.**

**a**, Structure of the  $\sigma_2$  receptor bound to PB28. Amino and carboxy termini are indicated. Membrane boundaries were calculated using the PPM server<sup>45</sup>. **b**, Cross-section of the  $\sigma_2$  receptor-binding pocket (left) and view of the entrance to the binding pocket from the membrane (right). **c**, View of the PB28-binding

pose, showing charge–charge interaction with Asp29 (black dotted line) and contacts with other binding pocket residues. **d**, Analogous structure of the roluperidone-binding pose. **e**, Structure of the  $\sigma_1$  receptor bound to PD144418 (PDB ID: SHK1). Amino acids that have similar roles and that are positioned in a similar orientation to amino acids in the  $\sigma_2$  receptor are indicated.

a biochemical and structural approach combined with computational docking to address these issues.

## Structure determination

The human  $\sigma_2$  receptor was expressed in *Sf9* insect cells, extracted with detergent and purified<sup>14</sup>. Size-exclusion chromatography with multi-angle light scattering (SEC–MALS) showed that the receptor is a dimer in solution. Notably, all members of the EXPERA family are either dimers or pseudo-dimers, although the functional role of dimerization remains unknown. Unlike the  $\sigma_1$  receptor, which can change oligomeric state in response to ligand binding<sup>24</sup>, the presence of ligands did not perturb the oligomeric state of  $\sigma_2$  (Extended Data Fig. 1a). As the human  $\sigma_2$  receptor was not tractable in structural studies, further experiments were performed with the homologous bovine  $\sigma_2$  receptor (Extended Data Fig. 1b). Circular dichroism experiments showed that the bovine  $\sigma_2$  receptor is 74% helical (Extended Data Fig. 1c). Thermal unfolding demonstrated that the receptor is highly stable, with a midpoint of the unfolding transition ( $T_m$ ) of 54 °C (Extended Data Fig. 1d). Crystals of the  $\sigma_2$  receptor were grown by the lipidic cubic phase method<sup>25</sup> (Extended Data Fig. 1e–g). Three datasets were collected for the receptor in complex with PB28<sup>5</sup>, roluperidone<sup>2</sup> and a ligand tentatively modelled as cholesterol (Extended Data Table 1). Molecular replacement was performed using a model derived from the structure of EBP<sup>26</sup> (see Methods).

## Overall structure of the $\sigma_2$ receptor

The three  $\sigma_2$  receptor crystal structures are similar, with a backbone root mean square deviation (RMSD) of 0.75 Å. As anticipated from SEC–MALS, the structures showed that  $\sigma_2$  is an intimately associated homodimer, burying 890 Å<sup>2</sup> of surface area in a dimer interface mainly formed by transmembrane helix 3 (TM3; Fig. 1a). The two protomers adopt the same conformation (backbone RMSD of 0.34 Å, 160 residues), with each adopting the expected four-helix bundle fold.

The four transmembrane helices of the protein are all kinked owing to the presence of proline residues in each, creating a ligand-binding cavity near the centre of the protein. This cavity is entirely occluded from solvent by extracellular loops 1 and 2, which form a well-ordered cap over the luminal surface of the protein. Asp56, which is crucial for ligand binding<sup>14</sup>, bridges extracellular loop 1 to TM4 using a hydrogen-bond

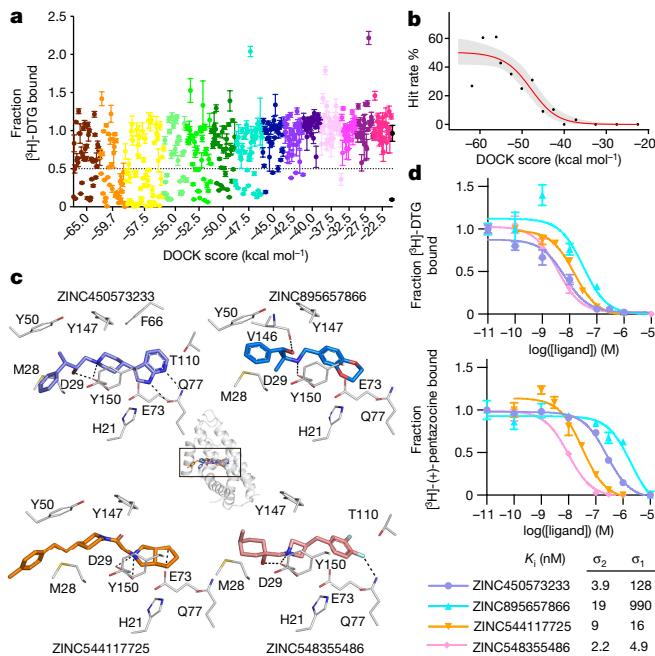
network (Extended Data Fig. 1h). Hence, Asp56 is likely to be important for receptor folding and not directly for ligand recognition<sup>14</sup>. Rather than opening to the ER lumen, the pocket opens laterally into the lipid bilayer (Fig. 1b), reminiscent of lipid-binding G-protein-coupled receptors (GPCRs)<sup>27</sup>, and its opening is lined with hydrophobic and aromatic residues. Ligands may enter through this opening in their neutral, deprotonated form and then become protonated in the binding site, forming a salt bridge with the conserved Asp29 (Fig. 1c, d). A second conserved acidic residue, Glu73, is located 3 Å away from Asp29, suggesting that these residues are hydrogen-bonded to each other, with Glu73 probably protonated.

The two  $\sigma$  receptors are not homologues and do not share the same fold; the  $\sigma_2$  receptor is a four-helix bundle, whereas the  $\sigma_1$  receptor has a  $\beta$ -barrel cupin fold<sup>28</sup>. Nevertheless, the binding pockets of the two receptors are similar (Fig. 1c–e), placing functionally comparable amino acids in cognate spatial positions, which is perhaps the result of convergent evolution and explains how two very different folds can share closely overlapping ligand recognition profiles.

Both  $\sigma$  receptors are homologues of proteins that catalyse the same step in sterol biosynthesis. The  $\sigma_1$  receptor is a homologue of ERG2, the fungal  $\Delta 8,7$  sterol isomerase; the  $\sigma_2$  receptor is a homologue of EBP, the mammalian  $\Delta 8,7$  sterol isomerase. Both EBP and ERG2 rely on two interacting acidic residues in their active site for catalysis, which occurs by protonation of the substrate at carbon 9 (C9) followed by proton abstraction from C7, shifting the double bond into the C8–C7 position. All necessary components for catalysis appear to be present in the  $\sigma_2$  receptor, yet it does not catalyse sterol isomerization. It can function neither *in vivo* to rescue a strain of yeast that lacks ERG2 (Extended Data Fig. 1k), nor *in vitro* to convert zymostenol to lathosterol (Extended Data Fig. 1l). The same is true for the  $\sigma_1$  receptor, which also has all the conserved residues required for catalysis but cannot rescue yeast that lack a sterol isomerase<sup>12</sup> (Extended Data Fig. 1k). It was recently reported that  $\Delta 8,9$  sterols can serve as signalling molecules<sup>29</sup>, which may hint at a possible physiological function of the  $\sigma$  receptors as sensors of these molecules evolved from enzymes that would modify them.

## Docking against the $\sigma_2$ receptor

Docking against the  $\sigma_2$  receptor had two goals: discovering novel and  $\sigma_2$ -selective chemotypes; and investigating how docking scores predict



**Fig. 2 | Docking 490 million molecules against the  $\sigma_2$  receptor.**

**a**, Displacement of the radioligand [ $^3\text{H}$ ]DTG by each of the 484 molecules tested at 1  $\mu\text{M}$  (mean  $\pm$  s.e.m. of three technical replicates). The molecules are coloured and grouped by docking score. Dashed line indicates 50% radioligand displacement. Dots below the dashed line represent confirmed binders, the numbers of which diminish with worsening docking score. **b**, Hit rate of 484 experimentally tested compounds plotted against docking energy. The docking score ( $\text{dock}_{50}$ ) and slope at the maximum ( $\text{slope}_{50}$ ) are  $-48 \text{ kcal mol}^{-1}$  and  $-4.2\%$  per  $\text{kcal mol}^{-1}$ , respectively. The grey band represents the 95% credible interval. **c**, Docked poses of four representative ligands from different scaffolds. **d**, Competition binding curves of the four molecules in **c** against the  $\sigma_2$  receptor (top) and against the  $\sigma_1$  receptor (bottom). Data are mean  $\pm$  s.e.m. from three technical replicates.

binding likelihood. This has to our knowledge been done only once before at scale, against the dopamine receptor, which revealed a sigmoidal relationship between hit rate (active ligands/number tested) and score<sup>30</sup>. The promiscuous  $\sigma_2$  site promised a higher hit rate, increasing the dynamic range of any relationship observed. Guided by score alone for most molecules picked, supplemented by manual selection among the highest-ranking docked molecules, we prioritized 577 molecules for synthesis, spread among 14 scoring bins, of which 484 compounds were successfully produced. We tested compounds at 1  $\mu\text{M}$  and defined as ‘hits’ those that displaced greater than 50% of the binding of [ $^3\text{H}$ ]ditolylguanidine ( $[^3\text{H}]$ DTG) to  $\sigma_2$ . A total of 127 out of 484 molecules qualified, accounting for 26% of compounds over the full scoring range and a 60% hit rate among the top-ranked molecules (Fig. 2a). Hit rates fell monotonically with score, as with the dopamine receptor<sup>30</sup>, with a slope of  $-4.2\%$  per  $\text{kcal mol}^{-1}$  in the inflection region, with one exception (below). The curve dropped from a hit rate of 61% at a docking score of about  $-60 \text{ kcal mol}^{-1}$  to 0% at the four lowest-scoring bins ( $-40$  to  $-22.5 \text{ kcal mol}^{-1}$ ) (Fig. 2b, Supplementary Fig. 1).

The highest-scoring bin had a hit rate of 27%—much lower than the 61% hit rate that was observed in the second-best-scoring bin. This dip in the hit-rate curve highlights defects in the scoring function. Many of the molecules in the top-scoring bin had unexpectedly low desolvation penalties (Extended Data Fig. 2a, b). DOCK 3.7 pre-calculates these penalties from one conformation among hundreds docked, which is not necessarily the highest scoring conformation against a target. Indeed, recalculating ligand desolvation using the docked conformation for molecules tested against  $\sigma_2$  and dopamine receptors increased

desolvation penalties for molecules in the top-scoring bin, reducing their ranking and so suggesting a method to improve the scoring function (Extended Data Fig. 2d).

To supplement molecules prioritized by score alone, we picked a comparable number of high-ranking molecules by human inspection<sup>30,31</sup>. In the top three scoring bins (139 molecules) the human-prioritized hit rate (67%) was higher than that by docking score alone (33%) (Extended Data Fig. 2e, f), and the human-prioritized molecules reached higher affinities (Extended Data Fig. 2g, h). Broadly, these patterns reflect what was observed against the dopamine receptor.

Seeking selective probes for the  $\sigma_2$  receptor, we measured competition binding curves for 14 docking hits with high radioligand displacement at 1  $\mu\text{M}$ . Inhibition constant ( $K_i$ ) values ranged from 2.4 to 68 nM. In competition binding versus the  $\sigma_1$  receptor (Fig. 2d, Extended Data Table 2, Supplementary Table 1), several of these had substantial selectivity for  $\sigma_2$  over  $\sigma_1$ , including ZINC450573233 and ZINC895657866, which were 30- and 46-fold selective, respectively.

We sought to improve the affinities of three potent ligands, each representing a different scaffold (Extended Data Fig. 3a–c). Twenty thousand analogues identified in SmallWorld (<https://sw.docking.org/>) from a virtual library of 28 billion were docked into the  $\sigma_2$  site (Methods, Supplementary Table 1). Of these, 105 were synthesized and tested, which improved the affinity of each scaffold by 2- to 18-fold (Extended Data Fig. 3a–c, Supplementary Table 1); for two chemotypes,  $\sigma_2$  selectivity improved by 47- and more than 250-fold (Z1665845742 and Z4857158944, respectively).

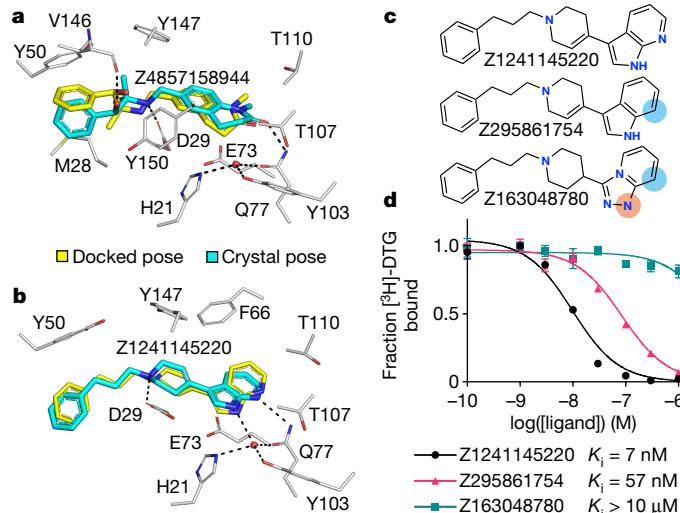
## Structures of $\sigma_2$ in complex with analogues

To test our docking poses, we determined the crystal structures of  $\sigma_2$  bound to two high-affinity ligands: Z1241145220 ( $\sigma_2 K_i = 3.7 \text{ nM}$ ; Protein Data Bank (PDB) ID: 7M95) and Z4857158944 ( $\sigma_2 K_i = 4 \text{ nM}$ ; PDB ID: 7M96). Electron density maps confirmed the docking predictions, with RMSD values between the crystallized and docked poses of 0.88 and 1.4  $\text{\AA}$ , respectively (Fig. 3a, b, Extended Data Table 1, Extended Data Fig. 1i). Newly predicted hydrogen-bond interactions with the backbone carbonyl of Val146, which were not seen in the roluperidone or PB28 complexes, corresponded well between docked and crystallographic poses. A hydrogen-bond interaction with Gln77 is also found in the roluperidone and cholesterol complexes (Fig. 1d, Extended Data Fig. 1j). The higher resolution of this structure, 2.4  $\text{\AA}$ , also revealed an ordered water molecule in one of the binding sub-sites, coordinated by residues His21, Tyr103 and Gln77, and by an azaindole nitrogen in Z1241145220 (Fig. 3b).

This water molecule was not modelled in the docked structure, so to investigate its role in ligand recognition we tested two analogues that were designed to disrupt the hydrogen bonds between Gln77 and the water (Fig. 3c). Z295861754, which should only hydrogen bond with the water but not with Gln77, exhibited an eightfold decrease in affinity, whereas Z163048780, which should not hydrogen bond with either Gln77 or the water, had a  $K_i$  value of greater than 10  $\mu\text{M}$  (Fig. 3d), indicating a crucial role of the water molecule for Z1241145220. We further generated a series of  $\sigma_2$  mutants in which the coordination of this water molecule was disrupted. Competition binding assays with Z1241145220 showed that mutating either His21 or Gln77 reduced the affinity by about 10-fold (Extended Data Fig. 3d–f). Together, these results show that the ordered water molecule is an integral part of the binding pocket and is required for high-affinity binding of Z1241145220, and probably other ligands.

## $\sigma_2$ ligands active in a mouse model of pain

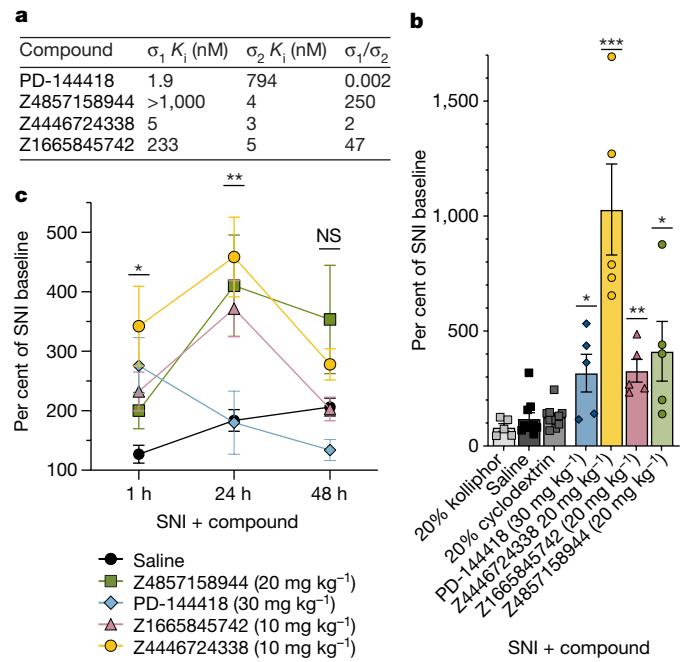
Genetic<sup>32,33</sup> and pharmacological<sup>34–36</sup> evidence suggests that the  $\sigma_1$  receptor has a role in chronic pain<sup>37</sup>. The discovery of the gene that encodes the  $\sigma_2$  receptor<sup>14</sup> made understanding and distinguishing



**Fig. 3 | High structural fidelity between docked and crystallographic poses of  $\sigma_2$  receptor ligands.** Ligand crystal poses (carbons in cyan) overlaid with respective docked poses (yellow).  $\sigma_2$  receptor carbons are in grey, oxygens in red, nitrogens in blue and sulfurs in yellow; hydrogen bonds are shown as black dashed lines. **a**, Z4857158944-bound complex (PDB ID: 7M96; RMSD = 1.4 Å). **b**, Z1241145220-bound complex (PDB ID: 7M95; RMSD = 0.88 Å). **c**, Two Z1241145220 analogues that disrupt the hydrogen bonds with Gln77 and the structural water. Blue and apricot circles depict differences between the analogues and the parent compound. **d**, Competition binding curve of compounds from **c**. The data are the mean  $\pm$  s.e.m. from three technical replicates.

the roles of  $\sigma_2$  and  $\sigma_1$  in this indication<sup>3,4</sup> fully possible. However, the limited availability of selective  $\sigma_2$  probes<sup>4</sup> hinders the ability to distinguish the effect of the two receptors. Accordingly, we treated mice with three high-affinity  $\sigma_2$  ligands with differing degrees of  $\sigma_2/\sigma_1$  selectivity: Z4857158944 (4 nM; more than 250-fold selective), Z1665845742 (5 nM; 47-fold selective), and Z4446724338 (3 nM; non-selective) (Fig. 4a). We also treated mice with PB28, a well-established 5 nM non-selective ligand<sup>8</sup>. In pharmacokinetics experiments, the three docking-derived ligands had substantial brain permeability, with brain-to-plasma ratios ranging from 3 to 16, and brain half-lives ranging from 1.2 to 12 h (Extended Data Table 3). PB28 also had high brain permeability and a relatively long half-life, although its maximum observed concentration ( $C_{max}$ ) in the brain was three- to eightfold lower than that of the new compounds. The high brain exposures of all four compounds encouraged us to examine them in a neuropathic pain model in mice.

We tested the efficacy of these ligands in the spared nerve injury (SNI) mouse model of neuropathic pain, in which two out of three branches of the sciatic nerve are transected<sup>9</sup>, resulting in mechanical hypersensitivity (allodynia) transmitted by the uninjured peripheral (sural) nerve. *In situ* hybridization of sections of dorsal root ganglia (DRG), where the cell bodies of sensory neurons that transmit the ‘pain’ message to the spinal cord reside, revealed the expression of both  $\sigma_1$  and  $\sigma_2$  receptors in many DRG neurons, including myelinated and unmyelinated subsets (Extended Data Fig. 4). The expression of  $\sigma_1$  or  $\sigma_2$  did not change in DRG neurons seven days after SNI. When administered systemically to SNI mice, both of the  $\sigma_2$ -selective ligands (Z1665845742 and Z4857158944) were anti-allodynic, increasing mechanical thresholds versus vehicle (Fig. 4b, Extended Data Fig. 4). This was comparable to the anti-allodynia conferred by a systemic injection of PD-144418, a  $\sigma_1$ -selective ligand. Notably, systemic injection of the non-selective ligand Z4446724338 dose-dependently increased the mechanical thresholds of SNI mice (Fig. 4b, Extended Data Fig. 4) with the highest dose completely reversing the SNI-induced mechanical allodynia (that is, thresholds returned to pre-injury levels)—a meaningfully higher level of anti-allodynia than



**Fig. 4 |  $\sigma_{1/2}$  ligands are anti-allodynic in a model of neuropathic pain.**

**a**, Selectivity of four ligands at  $\sigma_1$  and  $\sigma_2$ . PD-144418 values from the literature<sup>46</sup>. **b**, Response of mice to a von Frey filament after SNI. Ligands are compared to their vehicles (PD-144418 30 mg kg<sup>-1</sup> ( $n$  = 5) versus kolliphor ( $n$  = 5), one-way ANOVA,  $F(2,12)$  = 7.49,  $P$  = 0.008; Z4446724338 20 mg kg<sup>-1</sup> ( $n$  = 5) versus cyclodextrin ( $n$  = 10), one-way ANOVA,  $F(2,22)$  = 25.12,  $P$  = 0.00000021; Z4857158944 20 mg kg<sup>-1</sup> ( $n$  = 5) versus cyclodextrin ( $n$  = 10), one-way ANOVA,  $F(2,17)$  = 5.10,  $P$  = 0.02; Z1665845742 20 mg kg<sup>-1</sup> ( $n$  = 5) versus saline ( $n$  = 10), one-way ANOVA,  $F(3,31)$  = 6.18,  $P$  = 0.002; asterisks define individual group differences to respective vehicle control using Dunnett’s multiple comparisons post-hoc test; kolliphor versus PD-144418 30 mg kg<sup>-1</sup> ( $P$  = 0.009); cyclodextrin versus Z4446724338 20 mg kg<sup>-1</sup> ( $P$  < 0.001); cyclodextrin versus Z4857158944 20 mg kg<sup>-1</sup> ( $P$  = 0.01); saline versus Z1665845742 20 mg kg<sup>-1</sup> ( $P$  = 0.002); \* $P$  < 0.05, \*\* $P$  < 0.01, \*\*\* $P$  < 0.001. Data are mean  $\pm$  s.e.m. Also see Extended Data Fig. 4a. **c**, The anti-allodynic effects of  $\sigma_2$ , but not  $\sigma_1$ , ligands peak at 24 h after injection (two-way ANOVA; time  $\times$  treatment interaction:  $F(8,80)$  = 2.25,  $P$  = 0.03; time:  $F(2,76)$  = 5.09,  $P$  = 0.009; treatment:  $F(4,40)$  = 5.40,  $P$  = 0.001; four treatment groups ( $n$  = 10) except PD-144418 ( $n$  = 5); asterisks define difference between Z4446724338 and saline at 1 h ( $P$  = 0.03), 24 h ( $P$  = 0.008) and 48 h ( $P$  = 0.11) for simplicity; NS, not significant, \* $P$  < 0.05, \*\* $P$  < 0.01. Data are mean  $\pm$  s.e.m.

was observed with the selective  $\sigma_2$  ligands. Conversely, systemic injections of the non-selective PB28 ligand<sup>8</sup> produced mixed results, with anti-allodynic effects observed only in 60% of the mice (Extended Data Fig. 4). The much stronger anti-allodynia of Z4446724338 versus PB28 may reflect the substantially higher brain permeability of the former (Extended Data Table 3). None of the new  $\sigma_1$  and  $\sigma_2$  ligands were sedative on the rotarod test (Extended Data Fig. 4), indicating that their anti-allodynic effect was not due to motor impairment.

The anti-allodynia of the  $\sigma_2$ -selective ligands Z1665845742 and Z4857158944 suggests that this receptor is a potential target for managing neuropathic pain. However, because  $\sigma_2$  ligands are notoriously promiscuous, especially against GPCRs<sup>38,39</sup>, we counter-screened the three docking-derived ligands against potential off-targets. In a TANGO screen<sup>40</sup> of 320 GPCRs, the molecules did not act as agonists or inverse agonists against most targets (Extended Data Fig. 5a–c), and the few cases in which activity was observed did not replicate in concentration–response assays (Extended Data Fig. 5d–f, Supplementary Figs. 2, 3). We also did not observe substantial activity at the  $\mu$ -opioid receptor—a key pain target—in a G protein assay (Extended Data Fig. 5d). We further screened the compounds in binding assays against a panel of

19 targets including GPCRs, ion channels and transporters; no binding was observed for any pain-related targets (Supplementary Table 2). These observations suggest that the primary mechanism of action of these ligands is through the  $\sigma_2$  receptor. The stronger activity of the  $\sigma_1$  and  $\sigma_2$  crossreactive ligand Z4446724338 suggests that  $\sigma_{1/2}$  polypharmacology may further increase anti-allodynia.

### The effects of $\sigma_2$ ligands peak after 24 hours

In previous studies,  $\sigma_{1/2}$  ligands showed peak anti-allodynia up to 48 h after dosing<sup>3</sup>. This unusual behaviour was observed with ligands with mid-nanomolar potency and 9 to 14-fold selectivity versus the  $\sigma_1$  receptor. We further examined this with the selective ligands Z4857158944 and Z1665845742, and the non-selective ligand Z4446724338. The molecules were tested after SNI, at 1, 24 and 48 h after dosing. Consistent with previous reports, the anti-allodynia of the three new  $\sigma$  ligands increased over time, peaking at 24 h after injection (Fig. 4c. Extended Data Fig. 6). By contrast, the anti-allodynia of the selective  $\sigma_1$  ligand PD-144418 was not sustained 24 h or 48 h after injection. Furthermore, although the  $\sigma_2$ -selective compounds exhibited reduced anti-allodynia efficacy at early time points versus the non-selective ligand Z4446724338, all three compounds conferred similar antinociception by 24 h. This long-term activity cannot be easily explained by pharmacokinetics, as the brain half-life of all three compounds suggests minimal exposure past 12 h (Extended Data Table 3). Rather, this time course may reflect longer-term signalling or regulatory effects<sup>3</sup>.

To investigate tolerance, we also examined the effects of repeated injections of two of the lead compounds, Z4446724338 and Z4857158944. The antinociceptive effect of Z4446724338 persisted for the first three test days, and decreased slightly on the fourth day (Extended Data Figs. 4c, d, 6c, d). More tolerance was observed for compound Z4857158944; by the third injection, the antinociceptive effect was lost. Together, these results suggest that polypharmacology at the  $\sigma_1$  and  $\sigma_2$  receptor underlies an enhanced antinociceptive effect compared to selectivity for the  $\sigma_2$  receptor.

## Discussion

The structure, function, and biology of the  $\sigma_2$  receptor have remained enigmatic for 30 years. Its involvement in diverse biological processes and the lack of molecular data available have clouded its biological role. Four key observations from this study begin to shed light on these issues. First, high-resolution crystal structures of the  $\sigma_2$  receptor in complex with roluperidone and with PB28 reveal a ligand-binding site deeply embedded in the membrane (Fig. 1a, b), suggesting the possibility of a lipid as an endogenous ligand. The evolutionary connection of  $\sigma_2$  to EBP and the structure of the receptor bound to cholesterol support an ability to recognize sterols. The structures explain the simple pharmacophore of  $\sigma_2$  ligands—a cationic amine that ion-pairs with Asp29, while flanking hydrophobic and aromatic moieties are recognized by nearby aromatic residues. The structures also identify nearby polar residues, Gln77 and Thr110, which may aid in recognizing the hydroxyl moiety of sterols. These residues are rarely exploited by classic  $\sigma_2$  ligands but may provide new selectivity determinants for ligand discovery (Fig. 1c, d, Extended Data Fig. 1j). Second, by testing 484 compounds across ranks from a library of 490 million docked, a quantitative relationship emerged between docking score and the likelihood of binding (Fig. 2). Crystal structures of docking-derived ligands confirmed the docking predictions (Fig. 3a, b). Third, among the top-ranking docking hits were 31 novel scaffolds with potent affinities ( $K_i < 100$  nM) (Extended Data Table 2). Optimization of two of these led to potent ligands with 47 to more than 250-fold selectivity for the  $\sigma_2$  over the  $\sigma_1$  receptor (Supplementary Table 1). Fourth, three potent new  $\sigma_2$  chemotypes were tested for efficacy in a mouse model for neuropathic pain. All three were antialloodynic (Fig. 4). The expression pattern of the receptor

and the activity of the  $\sigma_2$ -selective ligands confirm a contribution of this receptor in pain processing and suggest its potential relevance in pain management.

The  $\sigma_2$  and the  $\sigma_1$  receptors are promiscuous, both binding to cationic amphiphiles, which leads to receptor cross-reactivity. Although many selective  $\sigma_1$  ligands—such as PD-144418 and (+)-pentazocine—have been described, there are far fewer selective ligands<sup>4,41</sup> for the  $\sigma_2$  receptor. We sought to optimize for such selectivity<sup>22,42,43</sup> using structure-based analoguing, which ultimately led to two selective chemotypes. We combined one of these with a close analogue that is  $\sigma_2$ -inactive, affording a ‘probe pair’ (Z1665845742 and Z1665798906; available through the probe collection of Sigma-Millipore: SML3141 and SML3142, respectively) (Supplementary Fig. 8). Such pairs can interrogate the role of the  $\sigma_2$  receptor in indications for which it has been widely mooted, including cancer<sup>1,19</sup>, schizophrenia<sup>2</sup> and Niemann-Pick disease<sup>15,16</sup>, with the activity of the non-binding member controlling for inevitable off-targets.

The very promiscuity of the  $\sigma_2$  receptor makes it a good template to investigate how docking score predicts binding likelihood, something that has been investigated once before at scale with the dopamine receptor<sup>30</sup>. As in that previous study, a sigmoidal relationship between score and hit rate emerged, here with hit rates peaking at over 60% (Fig. 2b). Unlike the dopamine receptor, which suffered from a long hit-rate plateau among the top-ranking molecules,  $\sigma_2$  hit rates continued to rise with docking score through most of the curve. The exception was among a thin slice of the molecules with the very highest scores, in which hit rates actually dropped owing to a subset of molecules that ‘cheat’ the scoring function (Extended Data Fig. 2), affording us the ability to improve it.

After completion of this study, a model of the  $\sigma_2$  receptor was released as part of the AlphaFold protein structure prediction database<sup>44</sup>. This model closely resembles the crystal structures solved here, with an overall backbone RMSD of 0.5 Å (Supplementary Fig. 4a). Of note for ligand discovery, binding-site residues have an all-atom RMSD value of less than 2 Å (Supplementary Fig. 4b). Despite the high fidelity of the model to the experimental structure, the 484 new compounds from docking against the crystal structure scored relatively poorly against the AlphaFold model (Supplementary Fig. 4c), reflecting a slightly contracted pocket in the model. It may yet be true that other ligands could be found that fit the AlphaFold model well and bind to the receptor. To investigate this, new prospective docking experiments will be informative.

Certain caveats should be taken into consideration. Although our ultimate goal was to find  $\sigma_2$ -selective ligands, a spectrum of affinities and selectivities for both  $\sigma$  receptors emerged, reflecting the similarities of their pockets and their well-known overlapping pharmacology (Fig. 1c–e). The high hit rates and potencies found here reflect a site unusually well-suited to ligand binding, something that is unlikely to translate to other targets. Although the docking-predicted poses for Z4857158944 and Z1241145220 were confirmed crystallographically, the important water-bridging interaction for Z1241145220 was missed.

The key observations of this work should not be obscured by these caveats. The crystal structures of  $\sigma_2$  receptors reveal the basis of its molecular recognition, and template structure-based campaigns for ligand discovery. From such campaigns emerged a predictive correlation between docking rank and likelihood of binding, and potent and selective  $\sigma_2$  ligands that may be used to probe receptor biology.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-04175-x>.

- Waarde, A. V. et al. Potential applications for sigma receptor ligands in cancer diagnosis and therapy. *Biochim. Biophys. Acta* **1848**, 2703–2714 (2015).
- Harvey, P. D. et al. Effects of roluperidone (MIN-101) on two dimensions of the negative symptoms factor score: reduced emotional experience and reduced emotional expression. *Schizophr. Res.* **215**, 352–356 (2020).
- Sahn, J. J., Mejia, G. L., Ray, P. R., Martin, S. F. & Price, T. J. Sigma 2 receptor/Tmem97 agonists produce long lasting antineuropathic pain effects in mice. *ACS Chem. Neurosci.* **8**, 1801–1811 (2017).
- Intagliata, S. et al. Discovery of a highly selective sigma-2 receptor ligand, 1-(4-(6,7-dimethoxy-3,4-dihydroisoquinolin-2(1H)-yl)butyl)-3-methyl-1H-benzo[d]imidazol-2(3H)-one (CM398), with drug-like properties and antinociceptive effects in vivo. *AAPS J.* **22**, 94 (2020).
- Quadir, S. G. et al. The Sigma-2 receptor / transmembrane protein 97 ( $\sigma_2$ R/TMEM97) modulator JWV-1034 reduces heavy alcohol drinking and associated pain states in male mice. *Neuropharmacology* **184**, 108409 (2021).
- Grundman, M. et al. A phase 1 clinical trial of the sigma-2 receptor complex allosteric antagonist CT1812, a novel therapeutic candidate for Alzheimer's disease. *Alzheimers Dement.* **5**, 20–26 (2018).
- Riad, A. et al. Sigma-2 receptor/TMEM97 and PGRMC-1 increase the rate of internalization of LDL by LDL receptor through the formation of a ternary complex. *Sci. Rep.* **8**, 16845 (2018).
- Abate, C. et al. PB28, the sigma-1 and sigma-2 receptors modulator with potent anti-SARS-CoV-2 activity: a review about its pharmacological properties and structure affinity relationships. *Front. Pharmacol.* **11**, 589810 (2020).
- Shields, S. D., Eckert, W. A. & Basbaum, A. I. Spared nerve injury model of neuropathic pain in the mouse: a behavioral and anatomic analysis. *J. Pain* **4**, 465–470 (2003).
- Hellewell, S. B. et al. Rat liver and kidney contain high densities of  $\sigma_1$  and  $\sigma_2$  receptors: characterization by ligand binding and photoaffinity labeling. *Eur. J. Pharmacol.* **268**, 9–18 (1994).
- Hellewell, S. B. & Bowen, W. D. A sigma-like binding site in rat pheochromocytoma (PC12) cells: decreased affinity for (+)-benzomorphans and lower molecular weight suggest a different sigma receptor form from that of guinea pig brain. *Brain Res.* **527**, 244–253 (1990).
- Hanner, M. et al. Purification, molecular cloning, and expression of the mammalian sigma1-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8072–8077 (1996).
- Langa, F. et al. Generation and phenotypic analysis of sigma receptor type I (sigma1) knockout mice. *Eur. J. Neurosci.* **18**, 2188–2196 (2003).
- Alon, A. et al. Identification of the gene that codes for the  $\sigma_2$  receptor. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7160–7165 (2017).
- Ebrahimi-Fakhari, D. et al. Reduction of TMEM97 increases NPC1 protein levels and restores cholesterol trafficking in Niemann-pick type C1 disease cells. *Hum. Mol. Genet.* **25**, 3588–3599 (2016).
- Bartz, F. et al. Identification of cholesterol-regulating genes by targeted RNAi screening. *Cell Metab.* **10**, 75 (2009).
- Sanchez-Pulido, L. & Ponting, C. P. TM6SF2 and MAC30, new enzyme homologs in sterol metabolism and common metabolic disease. *Front. Genet.* **5**, 439 (2014).
- Mahdessian, H. et al. TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8918 (2014).
- Vilner, B. J., John, C. S. & Bowen, W. D. Sigma-1 and sigma-2 receptors are expressed in a wide variety of human and rodent tumor cell lines. *Cancer Res.* **55**, 408–413 (1995).
- Scott, L. L. et al. Small molecule modulators of  $\sigma_2$ R/Tmem97 reduce alcohol withdrawal-induced behaviors. *Neuropsychopharmacology* **43**, 1867–1875 (2018).
- Vázquez-Rosa, E. et al. Neuroprotective efficacy of a sigma 2 receptor/TMEM97 modulator (DKR-1677) after traumatic brain injury. *ACS Chem. Neurosci.* **10**, 1595–1602 (2019).
- Stein, R. M. et al. Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. *Nature* **579**, 609–614 (2020).
- Schuller, M. et al. Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking. *Sci. Adv.* **7**, eabf8711 (2021).
- Schmidt, H. R. & Kruse, A. C. The molecular function of  $\sigma$  receptors: past, present, and future. *Trends Pharmacol. Sci.* **40**, 636–654 (2019).
- Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009).
- Long, T. et al. Structural basis for human sterol isomerase in cholesterol biosynthesis and multidrug recognition. *Nat. Commun.* **10**, 2452 (2019).
- Audet, M. & Stevens, R. C. Emerging structural biology of lipid G protein-coupled receptors. *Protein Sci.* **28**, 292–304 (2019).
- Schmidt, H. R. et al. Crystal structure of the human  $\sigma_1$  receptor. *Nature* **532**, 527–530 (2016).
- Hubler, Z. et al. Accumulation of 8,9-unsaturated sterols drives oligodendrocyte formation and remyelination. *Nature* **560**, 372–376 (2018).
- Lyu, J. et al. Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 (2019).
- Fischer, A., Smieško, M., Sellner, M. & Lill, M. A. Decision making in structure-based drug discovery: visual inspection of docking results. *J. Med. Chem.* **64**, 2489–2500 (2021).
- Cendán, C. M., Pujalte, J. M., Portillo-Salido, E., Montoliu, L. & Baeyens, J. M. Formalin-induced pain is reduced in  $\sigma_1$  receptor knockout mice. *Eur. J. Pharmacol.* **511**, 73–74 (2005).
- Puente, B. D. L. et al. Sigma-1 receptors regulate activity-induced spinal sensitization and neuropathic pain after peripheral nerve injury. *Pain* **145**, 294–303 (2009).
- Cendán, C. M., Pujalte, J. M., Portillo-Salido, E. & Baeyens, J. M. Antinociceptive effects of haloperidol and its metabolites in the formalin test in mice. *Psychopharmacology* **182**, 485–493 (2005).
- Romero, L. et al. Pharmacological properties of S1RA, a new sigma-1 receptor antagonist that inhibits neuropathic pain and activity-induced spinal sensitization. *Br. J. Pharmacol.* **166**, 2289–2306 (2012).
- Bruna, J. et al. Efficacy of a novel sigma-1 receptor antagonist for oxaliplatin-induced neuropathy: a randomized, double-blind, placebo-controlled phase IIa clinical trial. *Neurotherapeutics* **15**, 178–189 (2018).
- Vela, J. M., Merlos, M. & Almansa, C. Investigational sigma-1 receptor antagonists for the treatment of pain. *Expert Opin. Invest. Drugs* **24**, 883–896 (2015).
- Kooistra, A. J. et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. *Nucleic Acids Res.* **49**, D335–D343 (2021).
- Hauser, A. S. et al. Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54 (2018).
- Kroese, W. K. et al. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).
- Nastasi, G. et al. S2RSLDB: a comprehensive manually curated, internet-accessible database of the sigma-2 receptor selective ligands. *J. Cheminformatics* **9**, 3 (2017).
- Huang, X.-P. et al. Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. *Nature* **527**, 477–483 (2015).
- Wang, S. et al. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature* **555**, 269–273 (2018).
- Juniper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
- Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
- Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2021

## Methods

### Protein expression and purification for crystallography

The bovine  $\sigma_2$  receptor was cloned into pVL1392 with an N-terminal human protein C epitope tag followed by a 3C protease cleavage site. The construct was truncated after residue 168 to exclude the ER localization signal for better expression and to facilitate crystallization. This receptor construct was expressed in *Sf9* insect cells (Expression Systems) using the BestBac baculovirus system (Expression Systems) according to the manufacturer's instructions. Infection was performed when cell density reached  $4 \times 10^6$  cells per millilitre. Cells were shaken at 27 °C for 60 h before collection by centrifugation. Cell pellets were stored at -80 °C until purification.

During all purification steps, ligands (PB28, roluperidone, Z1241145220 and Z4857158944) were present in all buffers at 1  $\mu$ M. For the cholesterol-bound structure the protein was purified in the presence of 1  $\mu$ M DTG. Cell paste was thawed and cells were disrupted by osmotic shock in 20 mM HEPES pH 8, 2 mM magnesium chloride, 1:100,000 (v:v) benzonase nuclease (Sigma Aldrich) and cOmplete EDTA-free Protease Inhibitor Cocktail (Roche). Lysed cells were centrifuged at 50,000g for 15 min. After centrifugation, the supernatant was discarded, and the membrane pellets were solubilized with a glass Dounce tissue homogenizer in 20 mM HEPES pH 8, 250 mM NaCl, 10% (v/v) glycerol, 1% (w/v) lauryl maltose neopentyl glycol (LMNG; Anatrace) and 0.1% (w/v) cholesterol hemisuccinate (CHS; Steraloids). Samples were stirred at 4 °C for 2 h and then non-solubilized material was removed by centrifugation at 50,000g for 30 min. The supernatant was supplemented with 2 mM calcium chloride and filtered by a glass microfibre filter (VWR). Samples were then loaded by gravity flow onto 5 ml anti-protein C antibody affinity resin. Resin was washed with 10 column volumes of 20 mM HEPES pH 8, 250 mM NaCl, 2 mM calcium chloride, 1% (v/v) glycerol, 0.1% (w/v) LMNG and 0.01% (w/v) CHS, and then with 10 column volumes of 20 mM HEPES pH 8, 250 mM NaCl, 2 mM calcium chloride, 0.1% (v/v) glycerol, 0.01% (w/v) LMNG and 0.001% (w/v) CHS. The receptor was eluted with buffer containing 20 mM HEPES pH 8, 250 mM NaCl, 5 mM EDTA, 0.1% (v/v) glycerol, 0.01% (w/v) LMNG, 0.001% (w/v) CHS and 0.2 mg ml<sup>-1</sup> protein C peptide, in 1-ml fractions. Peak fractions were pulled and 3C protease was added (1:100 w:w) and incubated with the receptor at 4 °C overnight. Next the receptor was purified by SEC on a Sephadex S200 column (Cytiva) in 20 mM HEPES pH 8, 250 mM NaCl, 0.1% glycerol, 0.01% LMNG and 0.001% CHS. Peak fractions were pulled, calcium chloride was added to 2 mM and the sample was re-applied on the anti-protein C resin to remove uncleaved receptor. The column was washed with 5 column volumes and flow-through and wash fractions were pulled, concentrated and re-applied on SEC. Peak fractions were pulled, concentrated to 50 mg ml<sup>-1</sup>, and aliquoted. Protein aliquots were flash-frozen in liquid nitrogen and stored at -80 °C until use. Purity was evaluated by SDS-PAGE.

### Crystallography and data collection

Purified  $\sigma_2$  receptor was reconstituted into lipidic cubic phase (LCP) by mixing with a 10:1 (w:w) mix of monoolein (Hampton Research) with cholesterol (Sigma Aldrich) at a ratio of 1.5:1.0 lipid:protein by mass, using the coupled syringe reconstitution method<sup>25</sup>. All samples were mixed at least 100 times. The resulting phase was dispensed in 30–40-nl drops onto a hanging drop cover and overlaid with 800 nl of precipitant solution using a Gryphon LCP robot (Art Robbins Instruments). The PB28-bound crystals grew in 20–30% PEG 300, 0.1 M MES pH 6 and 600 mM NaCl. The roluperidone-bound crystals grew in 20% PEG 300, 0.1 M MES pH 6, 500 mM NaCl and 60 mM succinate. The Z1241145220-bound crystals grew in 30% PEG 300, 0.1 M MES pH 6 and 210 mM ammonium phosphate. The Z4857158944-bound crystals grew in 30% PEG 300, 0.1 M MES pH 6 and 560 mM ammonium phosphate. The cholesterol-bound crystals grew in 25% PEG 300, 0.1 M MES pH 6,

400 mM sodium citrate and 1% 1,2,3-heptanetriol. All crystals grew in the presence of 1  $\mu$ M ligand, except for the cholesterol structure, which had no ligand present during crystal growth. Crystals were collected using either MicroLoops LD or mesh loops (MiTeGen) and stored in liquid nitrogen until data collection. Data collection was performed at Advanced Photon Source GM/CA beamlines 23ID-B and 23ID-D. Data collection used a 10- $\mu$ m beam and diffraction images were collected in 0.2° oscillations at a wavelength of 1.254858 Å for the PB28-bound crystals and a wavelength of 1.033167 Å for all other crystals. A complete dataset was obtained from a single crystal in each case.

### Data reduction and refinement

Diffraction data were processed in HKL2000<sup>47</sup> and in XDS<sup>48</sup>, and statistics are summarized in Extended Data Table 1. The PB28-bound structure was solved using molecular replacement starting with a Rosetta<sup>49</sup> homology model generated using the structure of EBP (PDB ID: 6OHT). Matthews probability predicted four copies in the asymmetric unit. At first, a single copy of this model was placed using Phaser<sup>50</sup>, giving a marginally interpretable electron density map. This model did not fit well into density and was replaced with idealized helices that were used as a search model for an additional copy. The resulting dimer was duplicated and manually placed into unmodelled density. The resulting structure was iteratively refined in PHENIX<sup>51</sup> and manually rebuilt in Coot<sup>52</sup>. Final refinement statistics are summarized in Extended Data Table 1. The PB28 structure was used as a model for molecular replacement for all other datasets. In the case of the structure modelled as cholesterol-bound, electron density for a sterol-shaped ligand was observed (Extended Data Fig. 1i) and tentatively modelled as cholesterol on the basis of the high (millimolar) concentration of cholesterol in the crystallization conditions and the compatibility of cholesterol with the shape of the electron density in the binding pocket. The receptor was purified in the presence of ditolylguanidine (DTG), but no DTG was present in the precipitating solution, and electron density was clearly incompatible with bound DTG. We cannot exclude the possibility that some other compound structurally similar to cholesterol was carried through the purification and is the ligand observed in the binding pocket. Figures containing electron density or structures were prepared in PyMOL<sup>53</sup> v2.5 or UCSF Chimera<sup>54</sup> v1.15.

### Preparation of membranes for radioligand binding

The human  $\sigma_2$  receptor was cloned into a pcDNA3.1 (Invitrogen) mammalian expression vector with an amino-terminal protein C tag followed by a 3C protease cleavage site. Mutations were introduced by site-directed mutagenesis using HiFi HotStart DNA Polymerase (Kapa Biosystems). Expi293 cells were transfected using FectoPRO (Polyplus-transfection) according to the manufacturer's instructions. Cells were collected by centrifugation and lysed by osmotic shock in a buffer containing 20 mM HEPES, pH 7.5, 2 mM MgCl<sub>2</sub>, 1:100,000 (v/v) benzonase nuclease (Sigma Aldrich) and cOmplete Mini EDTA-free protease-inhibitor tablets (Sigma-Aldrich). The lysates were homogenized with a glass Dounce tissue homogenizer and then centrifuged at 20,000g for 20 min. After centrifugation, the membranes were resuspended in 50 mM Tris, pH 8.0, divided into 100- $\mu$ l aliquots, flash-frozen in liquid nitrogen and stored at -80 °C until use.

### Saturation and competition binding in Expi293 membranes

Saturation binding was performed with a method similar to that of a previous study<sup>55</sup>. In brief, membrane samples from Expi293 cells (Thermo Fisher Scientific) expressing wild-type or mutant  $\sigma_2$  receptor, prepared as described above, were thawed, homogenized with a glass Dounce and diluted in 50 mM Tris, pH 8.0. Binding reactions were done in 100  $\mu$ l, with 50 mM Tris pH 8.0, [<sup>3</sup>H]DTG (PerkinElmer), and supplemented with 0.1% bovine serum albumin to minimize non-specific binding. To assay non-specific binding, equivalent reactions containing 10  $\mu$ M haloperidol were performed in parallel. Competition assays were performed in

a similar manner with 10 nM [<sup>3</sup>H]DTG and the indicated concentration of the competing ligand. Samples were shaken at 37 °C for 90 min. Afterwards, the reaction was terminated by massive dilution and filtration over a glass microfibre filter with a Brandel harvester. Filters were soaked with 0.3% polyethyleneimine for at least 30 min before use. Radioactivity was measured by liquid scintillation counting. Data analysis was done in GraphPad Prism 9.0, with  $K_i$  values calculated by Cheng-Prusoff correction using the experimentally measured probe dissociation constant.

### Circular dichroism

Far-UV circular dichroism spectra (185–260 nm) were measured with a JASCO J-815 (JASCO), with a Peltier temperature controller and single cuvette holder and Spectra Manager II software for data collection and analysis. Data were collected using a 1-mm path-length cuvette, bandwidth of 1 nm, data pitch of 0.5 nm, scanning speed of 50 nm min<sup>-1</sup>, continuous scanning mode, and with 5 accumulations. Protein concentration was 0.25 mg ml<sup>-1</sup> (10 μM) in 10 mM potassium phosphate pH 7.4 and 250 mM potassium fluoride. Ligands were at 12 μM. Melt curves were measured at 222 nm between temperatures 20–95 °C, bandwidth of 1 nm, and a ramp rate of 1 °C per min with 10 s wait time. Calculation of  $T_m$  was done in Spectra Manager II by finding the peak of the first derivative of the melt curves, calculated using the Savitzky-Golay filter.

### SEC–MALS

The oligomeric state of the  $\sigma_2$  receptor was assessed by SEC–MALS using a Wyatt Dawn Heleos II multi-angle light scattering detector and Optilab TrEX refractive index monitor with an Agilent isocratic HPLC system Infinity II 1260. Receptor was prepared as described above, but with no ligand added during purification. The ligand-free receptor was diluted to 1 mg ml<sup>-1</sup> in SEC–MALS buffer (0.01% LMNG, 20 mM HEPES pH 7.5 and 150 mM sodium chloride). Ligands were added to a final concentration of 1 μM and the sample was incubated with ligand for 2 h at room temperature (21 °C). Separation steps were performed in SEC–MALS buffer with a Tosoh G4SWxl column at a flow rate of 0.5 ml min<sup>-1</sup>. Data analysis used the Astra software package v.6.1.4.25 (Wyatt) using the protein conjugate method with a dn/dc value of 0.21 (ml g<sup>-1</sup>) for detergent and 0.185 (ml g<sup>-1</sup>) for protein.

### Molecular docking

The  $\sigma_2$  receptor bound to cholesterol (PDB ID: 7MFI) was used in the docking calculations. The structure was protonated at pH 7.0 by Epik and PROPKA in Maestro<sup>56</sup> (2019 release). On the basis of the mutagenesis data<sup>14</sup>, Glu73 was modelled as a neutral residue. AMBER united atom charges were assigned to the structure. To model the more realistic low protein dielectric boundary of this site, we embedded the receptor into a lipid bilayer to capture its native environment in ER membrane, then followed this with a 50-ns coarse-grained molecular dynamics simulation with a restricted receptor conformation. A more detailed protocol can be found on the DISI wiki page ([http://wiki.docking.org/index.php/Membrane\\_Modeling](http://wiki.docking.org/index.php/Membrane_Modeling)). The volume of the low dielectric and the desolvation volume was extended out 2.2 Å and 1.2 Å, respectively, from the surface of protein and modelled lipid bilayer using spheres calculated by SPHGEN. Energy grids were pre-generated with AMBER force fields using CHEMGRID for van der Waals potential<sup>57</sup>, QNIFIT<sup>58</sup> for Poisson–Boltzmann-based electrostatic potentials and SOLVMAP<sup>59</sup> for ligand desolvation.

The resulting docking set-up was evaluated for its ability to enrich known  $\sigma_2$  ligands over property-matched decoys. Decoys are unlikely to bind to the receptor because despite their similar physical properties to known ligands, they are topologically dissimilar. We extracted 10 known  $\sigma_2$  ligands from ChEMBL (<https://www.ebi.ac.uk/chembl/>) including PB28 and roluperidone, the crystallographic poses of which we report here. Five-hundred and forty-two property-matched decoys were generated by the DUDE-Z pipeline<sup>60</sup>. Docking performance was evaluated on

the basis of the ability to enrich the knowns over the decoys by docking rank, using log-adjusted area under the curve (AUC) values (logAUC). The docking set-up described above was able to achieve a high logAUC of 39 and to recover the crystal poses of PB28 and roluperidone with RMSD values of 0.93 and 0.77 Å, respectively. This docking set-up gave the best retrospective enrichment and pose reproduction among three ligand-bound  $\sigma_2$  structures (Supplementary Fig. 5). We also constructed an ‘extrema’ set<sup>60</sup> of 61,687 molecules using the DUDE-Z web server (<http://tldr.docking.org>) to ensure that molecules with extreme physical properties were not enriched. The docking set-up enriched close to 90% mono-cations among the top 1,000 ranking molecules. To check whether the limited amounts of knowns and property-matched decoys overtrained the docking parameters, the enrichment test was run using 574 additional  $\sigma_2$  ligands from S2RSLDB<sup>41</sup> (<http://www.researchhdsf.unict.it/S2RSLDB>) against the ‘extrema’ set. The resulting high logAUC of 41 demonstrated the docking set-up was still able to enrich knowns over decoys on a 112-fold-larger test set, indicating the favourable docking parameters for launching an ultra-large-scale docking campaign.

Four-hundred and ninety million cations from ZINC15 (<http://zinc15.docking.org>), characterized by similar physical properties to  $\sigma_{1/2}$  known ligands (for instance, with calculated octanol–water partition coefficients (cLogP) ≤ 5 and with 250 Da < molecular weight ≤ 400 Da)), were then docked against the  $\sigma_2$  ligand-binding site using DOCK 3.8. Of these, 469 million molecules were successfully docked. On average, 3,502 orientations were explored and for each orientation, 183 conformations were on average sampled. In total, more than 314 trillion complexes were sampled and scored. The total calculation time was 177,087 h, or 3.7 calendar days on a cluster of 2,000 cores.

The top-ranking 300,000 molecules were filtered for novelty using the ECFP4-based Tanimoto coefficient ( $T_c$ ) against 2,232  $\sigma_{1/2}$  ligands in ChEMBL (<https://www.ebi.ac.uk/chembl/>) and 574  $\sigma_2$  ligands from S2RSLDB (<http://www.researchhdsf.unict.it/S2RSLDB>). Molecules with  $T_c \geq 0.35$  were eliminated. The remaining 196,170 molecules were clustered by an ECFP4-based  $T_c$  of 0.5, resulting in 33,585 unique clusters. From the top 5,000 novel chemotypes, molecules with >2 kcal mol<sup>-1</sup> internal strains were filtered out using strain\_rescore.py in Macro-model (2019 release). After filtering for novelty and diversity, the docked poses of the best-scoring members of each chemotype were manually inspected for favourable and diversified interactions with the  $\sigma_2$  site, such as the salt bridge with Asp29, the hydrogen bond with His21 or Val146 and the π-π stacking with Tyr50 or Trp49. Ultimately, 86 compounds were chosen for testing, 79 of which were successfully synthesized.

### Prediction of the hit-rate curve

To guide the design of scoring bins for the hit-rate curve, 1,000 docked poses were sampled in bins every 2.5 kcal mol<sup>-1</sup> from the best score of -65 kcal mol<sup>-1</sup> up to -22.5 kcal mol<sup>-1</sup>. We chose this 2.5 kcal mol<sup>-1</sup> distance between the bins to span the range with enough points (bins) to define a potential hit-rate versus docking-score curve. At the top of what we expected to be the curve, we increased the bin sizes because the density of molecules at these very highest ranks was relatively low. Correspondingly, at the lowest scores we added several more bins, also at a larger spacing, to help us to get a robust lower baseline. The estimated hit rate was calculated by the number of sensible docked poses divided by 1,000. The criteria to define a sensible docked pose include: (1) no unsatisfied hydrogen bond donors; (2) fewer than 3 unsatisfied hydrogen acceptors; (3) forms a salt bridge with Asp29; (4) total torsion strain energy of less than 8 units; and (5) maximum strain energy per torsion angle of less than 3 units. The first three filters were implemented on the basis of LUNA (<https://github.com/keiserlab/LUNA>), which calculated all the intra- and interactions of a docked pose with the receptor, then hashed them into a binary fingerprint. The strain energy was calculated by an in-house population-based method<sup>61</sup>. On the basis of the shape of the estimated prior curve (Supplementary

Fig. 6), more scoring bins are selected in the higher estimated hit-rate region:  $-65$ ,  $-59.73$  and  $-57.5$  kcal mol $^{-1}$ . After that, every scoring bin was 2.5 kcal mol $^{-1}$  from each other until  $-37.5$ . The last four bins were 5 kcal mol $^{-1}$  from each other. A total of 13,000 molecules sampled from these 14 scoring bins were filtered by novelty and internal torsion strain as described above. The remaining 9,216 novel and non-strained molecules were clustered by the LUNA 1,024-length binary fingerprint of a  $T_c = 0.32$ , resulting in 6,681 clusters. The first 40 chemotypes were attempted to be purchased from each scoring bin. After evaluation of the synthesis availability from the vendors, 491 molecules were ordered (Supplementary Tables 1, 3).

### Fitting of the hit-rate curve

To fit the Bayesian hit-rate models we used Stan<sup>62</sup> (v.2.21.2) via BRMS<sup>63</sup> (v.2.14.4), with generic parameters: iter=4000 and cores=4. Here are the model-specific parameters. For both hit-picking prior and posterior sigmoid models: formula=bmrs::formula(hit - top \* inv\_logit(hill\*4/top\*(dock\_energy - dock50)), top + hill + dock<sub>50</sub> - 1, nl=TRUE), where hill is scaled by 4/top so it is the slope of the curve at the dock<sub>50</sub> irrespective of the value of top. For the prior sigmoid model, prior=c(brms::prior(normal(.5, .2), lb=0, ub=1, nlnpar="top"), brms::prior(normal(-50, 10), nlnpar="dock<sub>50</sub>"), brms::prior(normal(-1, .1), ub=-.001, nlnpar="hill")), inits=function(){list(top=as.array(.5), dock<sub>50</sub>=as.array(-50), hill=as.array(-.1))}, family=gaussian(). Updating the prior sigmoid model with the mean expected hit rate for each computationally analysed tranche yielded an estimate and 95% credible interval for the sigma parameter for the Gaussian response of 20 [15, 30], but did not significantly adjust the distributions for top, hill or dock<sub>50</sub> (Supplementary Fig. 7). Therefore, to estimate the posterior sigmoid model, we transferred the per-parameter prior distributions and initial values and used the family=beroulli("identity"). To compare models, we used the loo package to add the Pareto smoothed importance sampling leave-one-out (PSIS-LOO) and Bayesian version of the R<sup>2</sup><sup>64</sup> (loo\_R2) information criteria. Figures were generated using the tidybayes<sup>65</sup>, ggplot2<sup>66</sup>, and tidyverse<sup>67</sup> packages in R<sup>68</sup>.

### Analoguing within the make-on-demand library

Using four primary docking hits (ZINC450573233, ZINC533478938, ZINC548355486 and ZINC895657866) as queries in SmallWorld (<https://sw.docking.org/>) from the 28B make-on-demand library, a subset of Enamine REAL space, 20,005 analogues were selected by its default settings, then docked into the  $\sigma_2$  site for potential favourable interactions with His21, Tyr50, Gln77 and Val146.

### Make-on-demand synthesis

Seventy-nine molecules that were prioritized by human inspection were delivered within 7 weeks with a 93% fulfilment rate, and 412 molecules by docking score alone were delivered within 4 weeks with an 82% fulfilment rate after a single synthesis attempt (Supplementary Tables 1, 3, 4). Most of the make-on-demand molecules were derived from the Enamine REAL database (<https://enamine.net/compound-collections/real-compounds>). See Supplementary Information for synthesis procedure and characterization of compounds.

### Yeast isomerase complementation assay

The human  $\sigma_2$  receptor, the *Saccharomyces cerevisiae* ERG2 and the human EBP were subcloned into the URA3 shuttle vector p416GPD. The plasmids were transformed into the Erg2-deficient *Saccharomyces cerevisiae* strain Y17700 (BY4742; MAT $\alpha$ ; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YMR202w::kanMX4) (Euroscarf) by the lithium acetate/single-stranded carrier DNA/polyethylene glycol method. A single colony was picked from a URA-selective plate and grown in suspension. Yeast was diluted in sterile water in a fivefold serial dilution starting from an optical density of 0.1. Two microlitres of the yeast dilutions was spotted on a URA-selective plate either in the absence or in the presence of sub-inhibitory

concentrations of cycloheximide (50 ng ml $^{-1}$ ) and grown at 30 °C for 24–48 h before imaging.

### Sterol isomerization enzymatic assay

The human EBP and  $\sigma_2$  were cloned into a pcDNA3.1 (Invitrogen) mammalian expression vector with FLAG and protein C affinity tag, respectively. Proteins were purified as described for crystallography preparations, except that no ligand was present during purification. After SEC, proteins were flash-frozen in liquid nitrogen and kept at  $-80$  °C until use. Zymostenol (CAS 566-97-2) and lathosterol (CAS 80-99-9) were purchased from Avanti Polar Lipids. For each sterol, a 2 $\times$  solution was prepared by first dissolving DDM in isopropanol to 1% (w/v) and dissolving sterols in chloroform to a concentration of 1 mg ml $^{-1}$ , followed by transferring 500  $\mu$ M of the sterols to a new vial, evaporating under argon and dissolving with DDM in a 1:20 (w/w) detergent-to-sterol ratio and a final 0.2% detergent in HEPES buffered saline (HBS; 20 mM HEPES pH 7.5, 150 mM NaCl). Proteins were diluted in HBS to 5  $\mu$ M. Individual sterol standards were prepared by mixing each sterol 1:1 with HBS. A mixed sterol standard was prepared by mixing both sterols in a 1:1 ratio. For the enzymatic reactions, sterols were mixed in a 1:1 ratio with the protein sample to give a final protein concentration of 2.5  $\mu$ M, sterol concentration of 250  $\mu$ M and detergent concentration of 0.1%, in HBS. Reactions were incubated for 1 h at 37 °C and then diluted 1:10 in methanol and kept at  $-20$  °C until analysis by liquid chromatography–mass spectrometry (LC–MS). Samples were analysed on a QE-plus mass spectrometer coupled to an Ultimate 3000 LC (Thermo Fisher Scientific) in a method modified from a previous study<sup>69</sup>. Five microlitres was injected on a Force PFPP column coupled with an Allure PFPP column (both 2 mm  $\times$  150 mm, Restek) maintained at 40 °C. The mobile phases were A: methanol:isopropyl alcohol:water:formic acid (80:10:10:0.02) 5 mM ammonium formate, and B: isopropyl alcohol. The gradient was as follows: 0% B for 15 min, then 100% B in 1 s, maintained at 100% B for 5 min, followed by 5 min re-equilibration at 0% B. The flow rate was 0.15 ml min $^{-1}$ . The mass spectrometer was acquiring in t-SIM mode for the [M – H<sub>2</sub>O + H] $^+$  ion (369.35158) with 70,000 resolution, and 0.5 m/z isolation. Standard samples for each compound were run first separately to obtain the retention time of each of the two isobaric compounds.

### $\mu$ OR activation assay

To measure  $\mu$ -opioid receptor ( $\mu$ OR) G<sub>10</sub>-mediated cAMP inhibition, 2.5 million HEK-293T cells (ATCC) were seeded in 10-cm plates. At 18 to 24 h later, after reaching 85–90% confluence, cells were transfected using a 1:3 ratio of human  $\mu$ OR and a split-luciferase-based cAMP biosensor (pGloSensorTM-22F; Promega). Transit 2020 (Mirus Biosciences) was used to complex the DNA at a ratio of 3  $\mu$ l Transit per  $\mu$ g DNA, in OptiMEM (Gibco-Thermo Fisher Scientific) at a concentration of 10 ng DNA per  $\mu$ l OptiMEM. Twenty-four hours later, cells were collected from the plate using Versene (PBS + 0.5 mM EDTA, pH 7.4) and plated in poly-D-lysine-coated white, clear-bottom 96-well assay plates (Corning Costar 3917) at a density of 35,000 cells per well and incubated at 37 °C with 5% CO<sub>2</sub> overnight. The next day, after aspiration of the culture medium, cells were incubated for 2 h covered, at room temperature, with 40  $\mu$ l assay buffer (CO<sub>2</sub>-independent medium, 10% FBS) supplemented with 2% (v/v) GloSensor reagent (Promega). To stimulate endogenous cAMP via  $\beta$ -adrenergic-G<sub>10</sub> activation, 5 $\times$  drugs were prepared in 10 $\times$  isoproterenol containing assay buffer (200 nM final concentration). For naloxone competition experiments, 5 $\times$  naloxone (1  $\mu$ M final concentration) was also added to each well. Luminescence was immediately quantified using a BMG Clariostar microplate reader. Data were analysed using nonlinear regression in GraphPad Prism 9.0 (Graphpad Software).

### Off-target counterscreens

Screening of compounds in the PRESTO-Tango GPCRome was accomplished as previously described<sup>39</sup> with several modifications. First,

# Article

HTLA cells were plated in DMEM with 10% FBS and 10 U ml<sup>-1</sup> penicillin-streptomycin. Next, the cells were transfected using an in-plate PEI method<sup>70</sup>. PRESTO-Tango receptor DNAs were resuspended in OptiMEM and hybridized with PEI before dilution and distribution into 384-well plates and subsequent addition to cells. After overnight incubation, drugs were added to cells at a 10  $\mu$ M final concentration without replacement of the medium. The remaining steps of the PRESTO-Tango protocol were followed as previously described. For those six receptors for which activity was reduced to less than 0.5-fold of basal levels of relative luminescence units or for the one receptor for which basal signalling was increased to greater than threefold of basal levels, assays were repeated as a full dose-response assay. Activity for none of the seven could be confirmed, and we discount the apparent activity seen in the single-point assay.

The radioligand binding screen of off-targets was performed by the National Institutes of Mental Health Psychoactive Drug Screen Program (PDSP)<sup>71</sup>. Detailed experimental protocols are available on the NIMH PDSP website at <https://pdsp.unc.edu/pdspweb/content/PDSP%20Protocols%20II%202013-03-28.pdf>.

## Cell lines

All cell lines in this study were not authenticated. All cells used in this study are commercial and were obtained from vendors as indicated. Cells were confirmed to be free of mycoplasma.

## Animals and ethical compliance

Animal experiments were approved by the UCSF Institutional Animal Care and Use Committee and were conducted in accordance with the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory animals. Adult (8–10 weeks old) male C56BL/6 mice (strain 664) were purchased from The Jackson Laboratory. Mice were housed in cages on a standard 12:12-h light-dark cycle with food and water ad libitum. We did not perform sample-size calculations. We modelled our sample sizes for behavioural studies on previous studies using a similar approach to our own, which have been demonstrated to be capable of detecting significant changes<sup>72,73</sup>. The mice were randomly assigned to the treatment group and control group. For behavioural experiments, mice were initially placed into one cage and allowed to free run for a few minutes. Next, each mouse was randomly picked up, injected with the drug or vehicle control, and placed into a separate cylinder before the behaviour test. All experiments were for animal behaviour and followed this randomization protocol. For all behavioural testing the experimenter was always blind to treatment. All experiments were in animals and under blinding conditions.

## Compounds

All ligands used in the animal studies were synthesized by Enamine (<https://enamine.net/>) (Supplementary Table 5) and dissolved 30 min before testing. PB28 and Z1665845742 were resuspended in 0.9% NaCl. Z4857158944 and Z4446724338 were resuspended in 20% cyclodextrin. PD-144418 was resuspended in 20% kolliphor.

## Behavioural analyses

For all behavioural tests, mice were first habituated for 1 h in Plexiglas cylinders. The experimenter was always blind to treatment. All tests were conducted 30 min after subcutaneous injection of the compounds. Hindpaw mechanical thresholds were determined with von Frey filaments using the up-down method<sup>74</sup>. For the ambulatory (rotarod) test, mice were first trained on an accelerating rotating rod, 3 times for 5 min, before testing with any compound.

## SNI model of neuropathic pain

Under isoflurane anaesthesia, two of the three branches of the sciatic nerve were ligated and transected distally, leaving the sural nerve intact. Behaviour was tested 7 to 14 days after injury and *in situ* hybridization was performed one week after injury.

## In situ hybridization

*In situ* hybridization was performed using fresh DRG tissue from adult mice (8–10 weeks old), following the protocol of Advanced Cell Diagnostics and as previously described<sup>75</sup>. All images were taken on an LSM 700 confocal microscope (Zeiss) and acquired with ZEN 2010 (Zeiss). Adjustment of brightness and contrast and changing of artificial colours (LUT) were done with Photoshop. The same imaging parameters and adjustments were used for all images within an experiment.

## Statistical analyses of animal studies

All statistical analyses of the animal studies were performed with GraphPad Prism 8.0 (GraphPad Software) unless otherwise noted. All data are reported as mean  $\pm$  s.e.m. unless otherwise noted. Dose-response experiments were analysed with a one-way ANOVA and time-course experiments were analysed with a two-way ANOVA, and both experiments used Dunnett's multiple comparison post-hoc test to determine differences between specific treatments and vehicle controls visualized in the figures. Rotarod experiments were analysed using a one-way ANOVA (saline, Z1665845742, and Z4857158944) or unpaired two-tailed Student's *t*-test (kolliphor and Z4446724338). Details of analyses, including the number of tested animals and groups, degrees of freedom and *P* values, can be found in the figure legends.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The coordinates and structure factors for PB28-bound  $\sigma_2$ , roluperidone-bound  $\sigma_2$ , Z1241145220-bound  $\sigma_2$ , Z4857158944-bound  $\sigma_2$  and cholesterol-bound  $\sigma_2$  have been deposited in the PDB with accession codes 7M93, 7M94, 7M95, 7M96 and 7MFI, respectively. The identities of the compounds docked in this study are freely available from the ZINC database (<http://zinc15.docking.org>) and active compounds may be purchased from Enamine. Any other data relating to this study are available from the corresponding authors on reasonable request. Source data are provided with this paper.

## Code availability

DOCK 3.7 is freely available for non-commercial research <http://dock.compbio.ucsf.edu/DOCK3.7/>. A web-based version is available at <http://blaster.docking.org/>.

48. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 132 (2010).
49. Alford, R. F. et al. An integrated framework advancing membrane protein modeling and design. *PLoS Comput. Biol.* **11**, e1004398 (2015).
50. Mccoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 674 (2007).
51. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
52. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
53. The PyMOL Molecular Graphics System v.2.5 (Schrödinger).
54. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
55. Chu, U. B. & Ruoho, A. E. Sigma receptor binding assays. *Curr. Protoc. Pharmacol.* **71**, 1.34.1–1.34.21 (2015).
56. Weiner, S. J. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
57. Meng, E. C., Shoichet, B. K. & Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **13**, 505–524 (1992).
58. Gallagher, K. & Sharp, K. Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* **75**, 769–776 (1998).
59. Mysinger, M. M. & Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50**, 1561–1573 (2010).
60. Stein, R. M. et al. Property-unmatched decoys in docking benchmarks. *J. Chem. Inf. Model.* **61**, 699–714 (2021).

61. Gu, S., Smith, M. S., Yang, Y., Irwin, J. J. & Shoichet, B. K. Ligand strain energy in large library docking. *J. Chem. Inf. Model.* **61**, 4331–4341 (2021).
62. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
63. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).
64. Gelman, A., Goodrich, B., Gabry, J. & Vehtari, A. R-squared for Bayesian regression models. *Am. Statistician* **73**, 307–309 (2018).
65. Kay, M. *tidybayes: tidy data and geoms for Bayesian models*. <https://doi.org/10.5281/zenodo.1308151> (Zenodo, 2020).
66. Wickham, H. *ggplot2: elegant graphics for data analysis* (Springer, 2016).
67. Wickham, H. et al. Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
68. R Core Team. *R: a Language and Environment for Statistical Computing*. <http://www.R-project.org> (R Foundation for Statistical Computing, 2018).
69. Skubic, C., Vovk, I., Rozman, D. & Križman, M. Simplified LC-MS method for analysis of sterols in biological samples. *Molecules* **25**, 4116 (2020).
70. Longo, P. A., Kavran, J. M., Kim, M.-S. & Leahy, D. J. Transient mammalian cell transfection with polyethylenimine (PEI). *Methods Enzymol.* **529**, 227–240 (2013).
71. Besnard, J. et al. Automated design of ligands to polypharmacological profiles. *Nature* **492**, 215–220 (2012).
72. Scherrer, G. et al. Dissociation of the opioid receptor mechanisms that control mechanical and heat pain. *Cell* **137**, 1148–1159 (2009).
73. Muralidharan, A. et al. Identification and characterization of novel candidate compounds targeting 6- and 7-transmembrane  $\mu$ -opioid receptor isoforms. *Br. J. Pharmacol.* **178**, 2709–2726 (2021).
74. Chaplan, S. R., Bach, F. W., Pogrel, J. W., Chung, J. M. & Yaksh, T. L. Quantitative assessment of tactile allodynia in the rat paw. *J. Neurosci. Methods* **53**, 55–63 (1994).
75. Solorzano, C. et al. Primary afferent and spinal cord expression of gastrin-releasing peptide: message, protein, and antibody concerns. *J. Neurosci.* **35**, 648–657 (2015).
76. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
77. Liebschner, D. et al. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D* **73**, 148–157 (2017).

**Acknowledgements** Funding to support this research was provided by NIH grant R01GM119185, the Vallee Foundation and the Sanofi iAwards program (A.C.K.); by DARPA grant HR0011-19-2-0020 and NIH grant R35GM122481 (B.K.S.); and by grant GM133836 (J.J.I.). GM/CA @ APS has been funded by the National Cancer Institute (ACB-12002) and the National

Institute of General Medical Sciences (AGM-12006 and P30GM138396). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. The Eiger 16M detector at GM/CA-XSD was funded by NIH grant S10 OD012289. We thank K. Arnett and the Harvard Center for Macromolecular Interactions for support with biophysical experiments including circular dichroism and SEC-MALS, and C. Vidoudez and the Harvard Center for Mass Spectrometry for performing mass spectrometry analysis of sterols. Molecular graphics and analyses were performed with UCSF Chimera, which was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

**Author contributions** A.A. performed cloning, mutagenesis, protein purification, SEC-MALS experiments, circular dichroism measurements, crystallization, X-ray data collection and processing, structure determination and refinement, radioligand binding, yeast complementation experiments and the sterol isomerization enzymatic assay. J.L. conducted the docking, chemoinformatics analyses, docking energy analysis and ligand picking, assisted in the latter by T.A.T. and B.K.S. J.M.B. conducted and analysed the mouse allodynia experiments with assistance from V.C., as well as the receptor expression experiments, which were supervised and co-analysed by A.I.B. M.J.O. conducted the Bayesian analysis of docking scores versus hit rates. C.M.W. tested molecules for activity against the  $\mu$ OR. X.-P.H. and Y.L. tested compounds against the GPCRome and other off-targets, with supervision from B.L.R. Y.S.M. supervised the synthesis of molecules from the virtual library. D.S.R. participated in the creation of Enamine REAL library. J.J.I. was responsible for the building of the version of the ZINC library that was docked. A.C.K., B.K.S. and A.I.B. supervised the project. The manuscript was written by A.A., J.L., B.K.S. and A.C.K. with input from other authors.

**Competing interests** A.C.K. is a founder and consultant for biotechnology companies Tectonic Therapeutic and Seismic Therapeutic, as well as the Institute for Protein Innovation, a non-profit research institute. B.K.S. is a founder of Epiodyne, a company active in analgesia, and of Blue Dolphin, which undertakes fee-for-service ligand-discovery.

**Additional information**

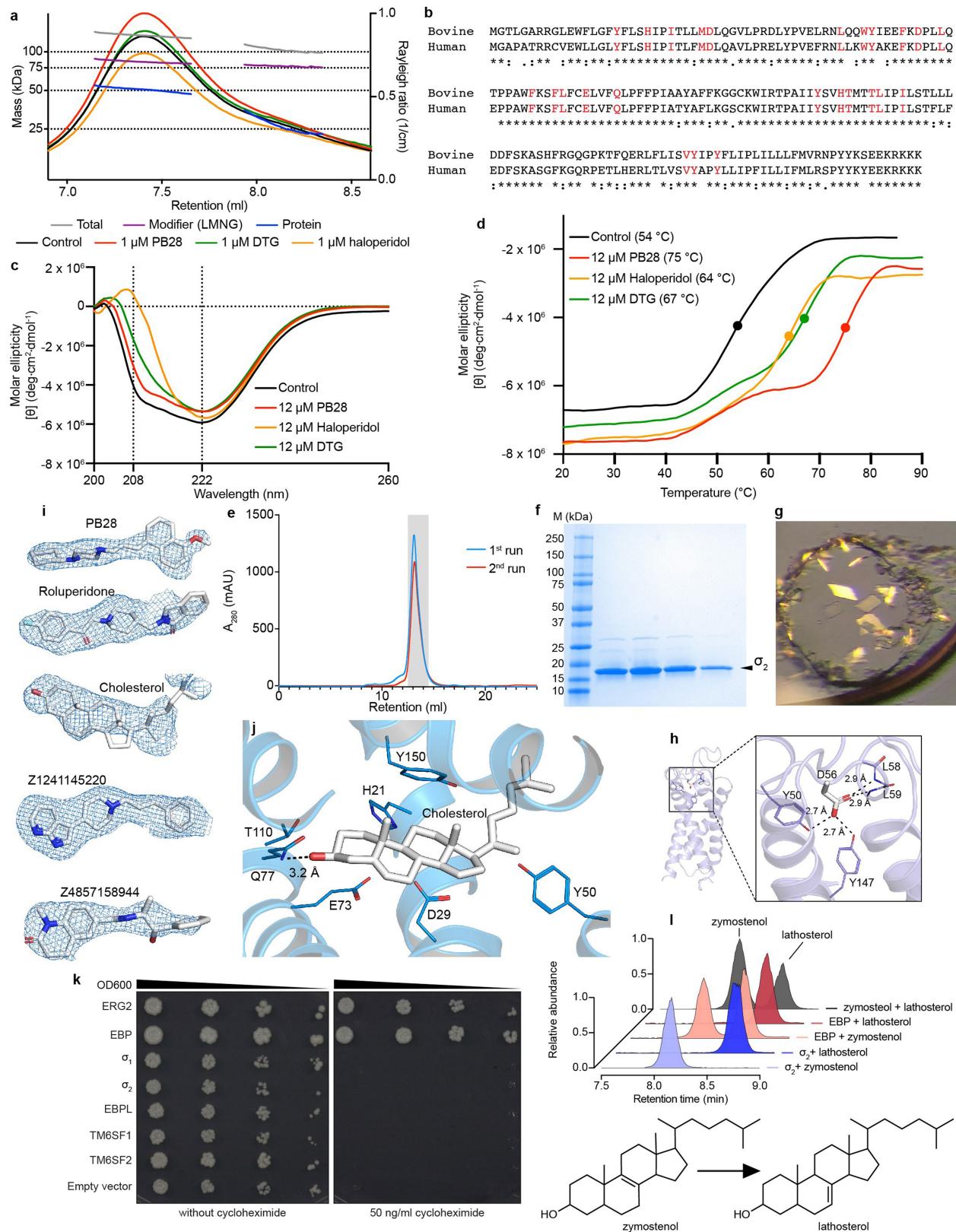
**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-04175-x>.

**Correspondence and requests for materials** should be addressed to Allan Basbaum, Brian K. Shoichet or Andrew C. Kruse.

**Peer review information** *Nature* thanks Kevin Jones and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

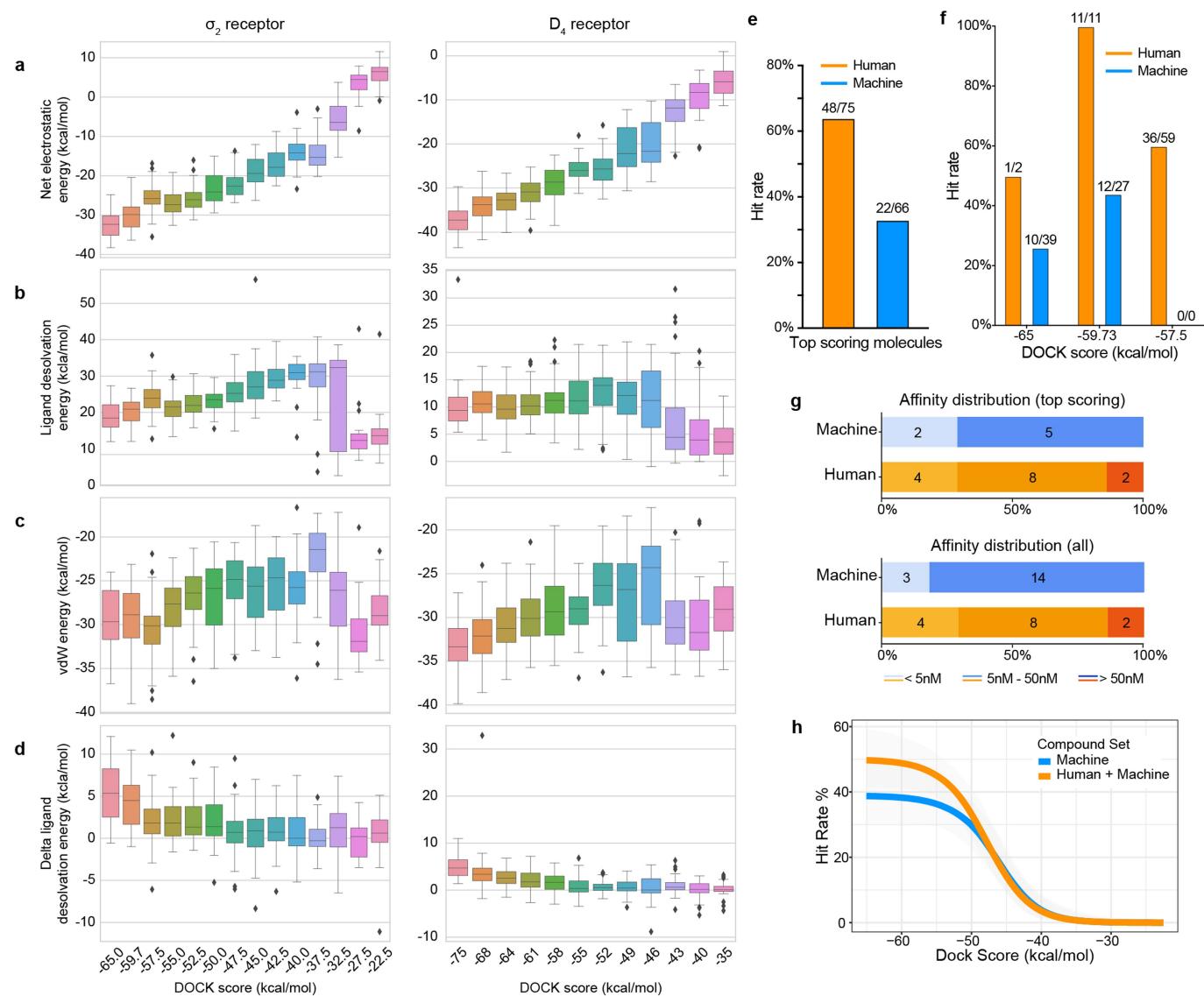
## Article



**Extended Data Fig. 1** | See next page for caption.

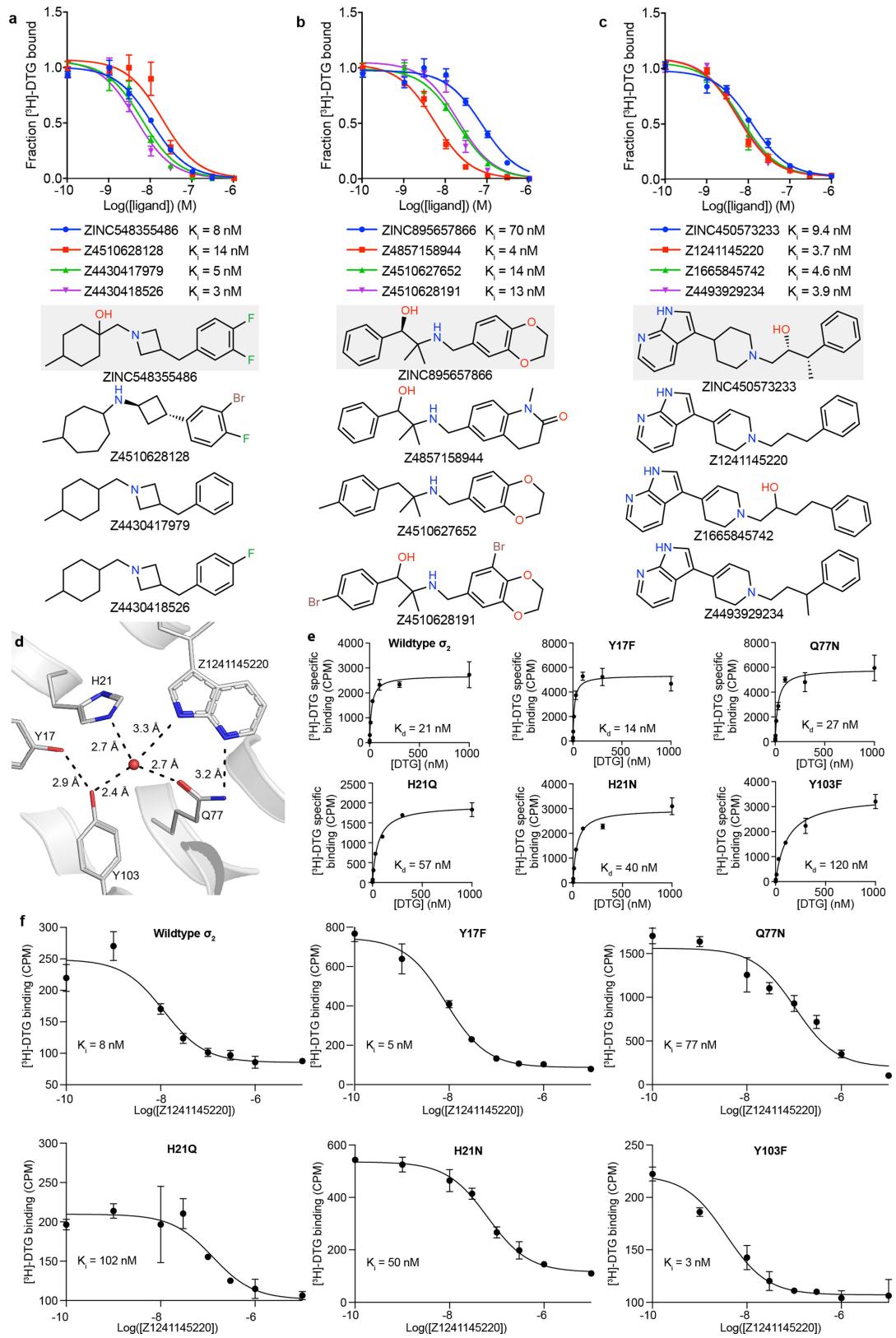
**Extended Data Fig. 1 | Characterization of the  $\sigma_2$  receptor.** **a**, SEC–MALS of the human  $\sigma_2$  receptor. The  $\sigma_2$  receptor was run either without ligand or with 1  $\mu$ M of the indicated ligand. Lines indicate calculated total mass (grey), detergent micelle (purple), and protein (blue). **b**, Sequence alignment between the human and bovine  $\sigma_2$  protein sequences performed using T-coffee<sup>76</sup>. Residues that line the binding pocket are marked in red. **c**, Circular dichroism analysis of the bovine  $\sigma_2$  receptor alone (black) or with the indicated ligand. Data is representative of multiple experiments. **d**, Circular dichroism melting curves of the bovine  $\sigma_2$  receptor. Temperature was raised from 20 °C to 90 °C and molar ellipticity was measured at 222 nm. Protein was incubated either with or without indicated ligand at 12  $\mu$ M. Melting temperatures for each measurement are indicated with a circle. Data is representative of multiple experiments **e**, SEC of the bovine  $\sigma_2$  receptor. Blue trace is after proteolytic tag removal. Red trace is protein applied on size exclusion after reapplying the tag-free protein on affinity resin to remove proteins with intact tags. The trace presented is representative of multiple purifications. **f**, Analysis of receptor purity after the second SEC using SDS–PAGE. Grey rectangle in **e** represents fractions chosen for analysis. The SDS-PAGE presented here is representative of multiple purifications. See Source Data for uncropped version. **g**, Crystals of bovine  $\sigma_2$  receptor in the lipidic cubic phase. **h**, Aspartate 56 (D56) is important

for receptor structure but not for ligand binding. A tight network of hydrogen bonds that bridges extracellular loop 1 to TM4 is depicted with black dashed lines. **i**, Electron density maps for the various ligands. Polder maps<sup>77</sup> were calculated in PHENIX. Maps are contoured at a level of 3  $\sigma$ . **j**, View of cholesterol-binding pose, showing contacts with other binding pocket residues. Hydrogen bonds are marked with black dashed lines. **k**, Yeast complementation assay. A  $\Delta$ ERG2 yeast strain was transformed with plasmids harbouring the indicated genes. Yeast cells were grown to logarithmic phase and diluted to OD600 of 0.1, and then further diluted in a five-fold serial dilution series. Two microliters of each dilution were spotted on plates. Yeast cells were grown either in permissive conditions of no cycloheximide or in the restrictive conditions of 50 ng/ml cycloheximide, which requires functional  $\Delta$ 8-9 sterol isomerase activity for viability. ERG2 and EBP can act as sterol isomerases and rescue the growth of  $\Delta$ ERG2 *S. cerevisiae* whereas the  $\sigma_2$  receptor, the  $\sigma_1$  receptor, or any other member of the EXPERA family cannot. **l**, EBP can catalyse the conversion of zymostenol to lathosterol whereas  $\sigma_2$  cannot. Standards are in dark grey. EBP converts zymostenol to lathosterol (apricot) but does not convert lathosterol to zymostenol (dark red). The  $\sigma_2$  receptor does not convert lathosterol to zymostenol (dark blue) or zymostenol to lathosterol (light purple). Structures of zymostenol and lathosterol are depicted below the traces.



**Extended Data Fig. 2 | Comparisons of the distribution of docking scores.** **a–d.** The distribution of docking scores of tested molecules for hit rate curves against  $\sigma_2$  (left column) and  $D_4$  (right column) receptors. All tested molecules are grouped based on docking score bins. The distributions are shown in box plots for **a**, net electrostatic energy, **b**, ligand desolvation energy, **c**, van der Waals (vdW) energy and **d**, delta ligand desolvation energy after recalculating atomic desolvation energy based on the docked pose. **e–h.** Comparison of hit rates and affinities achieved by combined docking score and human inspection and these achieved by docking score alone. **e**, Overall hit rates for selecting compounds from the first 3 scoring bins by each strategy: human prioritization and docking score (orange), or docking score alone (blue). Hit rate is the ratio of active compounds/tested compounds; the raw numbers appear at the top of each bar. **f**, Hit rates for selecting compounds at different scoring ranges by

each strategy: human prioritization and docking score (orange) or docking score alone (blue). **g**, Distribution of the binding affinity level among the hits from **e** (top panel). We measured competition binding curves for 14 docking hits from human prioritization and docking score, and 7 hits from the docking score alone. These are divided into three affinity ranges: < 5 nM; 5 nM–50 nM; > 50 nM; Distribution of the binding affinity level among the hits from all different scoring ranges (bottom panel). We measured competition binding curves for 14 docking hits from human prioritization and docking score, and 17 hits from the docking score alone. **h**, Hit-rate curve comparison with/without human picks. The hit rate without human picks at the top plateau is 39% and at the bottom plateau is 0%, and the docking score ( $\text{dock}_{50}$ ) and slope at the maximum ( $\text{slope}_{50}$ ) are  $-46.5 \text{ kcal mol}^{-1}$  and  $-3.5\%$  per  $\text{kcal mol}^{-1}$ , respectively.

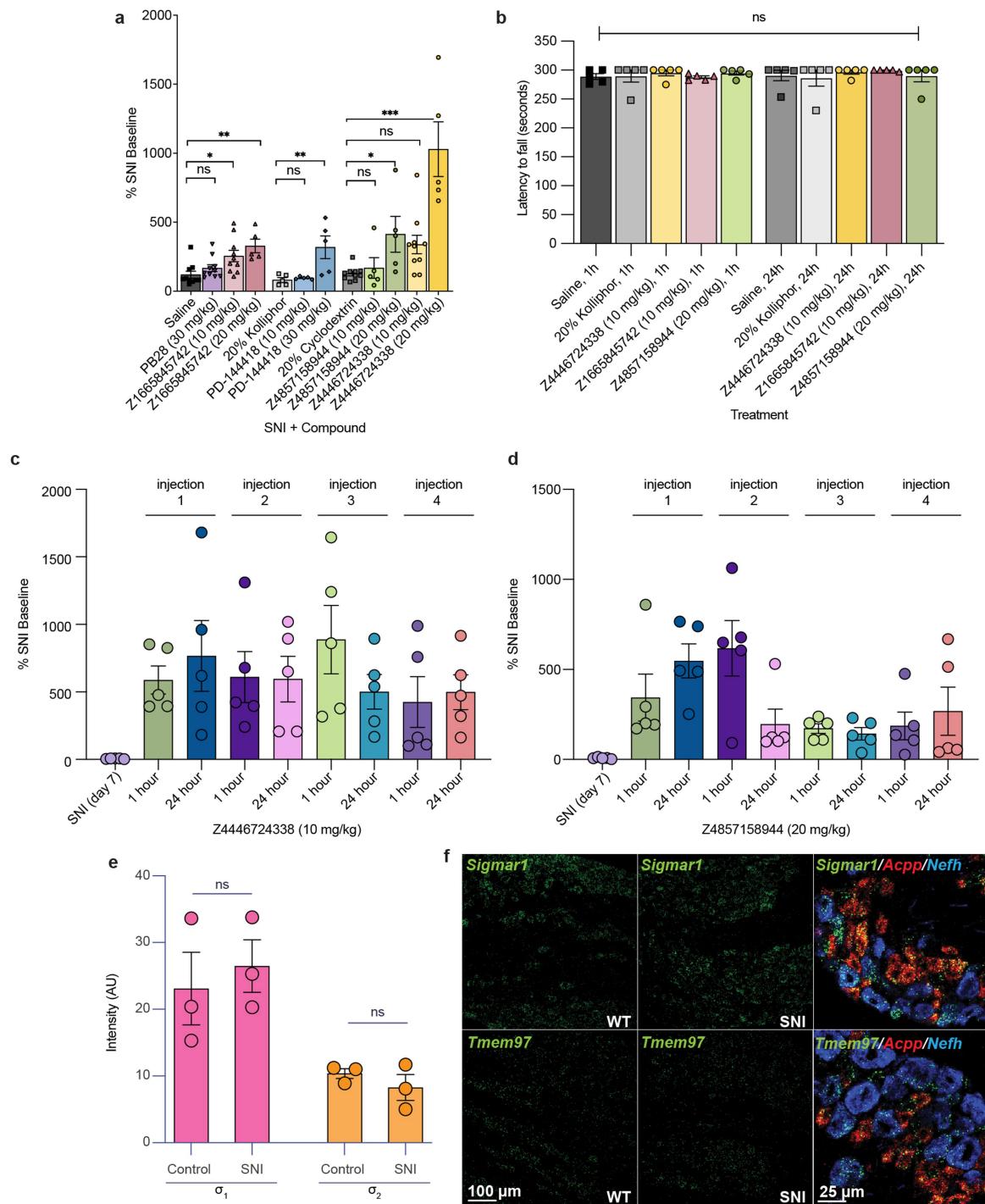


**Extended Data Fig. 3** | See next page for caption.

# Article

**Extended Data Fig. 3 | Analogues of  $\sigma_2$  receptor ligands and the effect of a structural water molecule.** **a–c**, Initial hits and selected analogues of  $\sigma_2$  receptor ligands. Competition binding curves on the top panel, 2D drawings of compounds are on the bottom panel. Parent compound is indicated by grey background. Points shown as mean  $\pm$  s.e.m. from three technical replicates. **a**, Parent compound ZINC548355486 and its three potent analogues. **b**, Parent compound ZINC895657866 and its three potent analogues. **c**, Parent compound ZINC450573233 and its three potent analogues. **d–f**, The binding site of the  $\sigma_2$  receptor contains a structural water. **d**, Water coordination at the

binding site of the  $\sigma_2$  receptor. Water molecule is depicted as a red sphere. Hydrogen bonds are indicated by black dashed lines. **e**, Saturation binding curve to measure the dissociation constant ( $K_d$ ) of [ $^3$ H]DTG for the various mutants of  $\sigma_2$  receptor meant to disrupt water coordination. Residues proximal to the structural water were chosen for mutation. Residues were mutated to the indicated amino acid. Points shown as mean  $\pm$  s.e.m. from three technical replicates. **f**, Competition binding measurement of affinity of Z1241145220 in various mutants of  $\sigma_2$ . Points shown as mean  $\pm$  s.e.m. from three technical replicates.



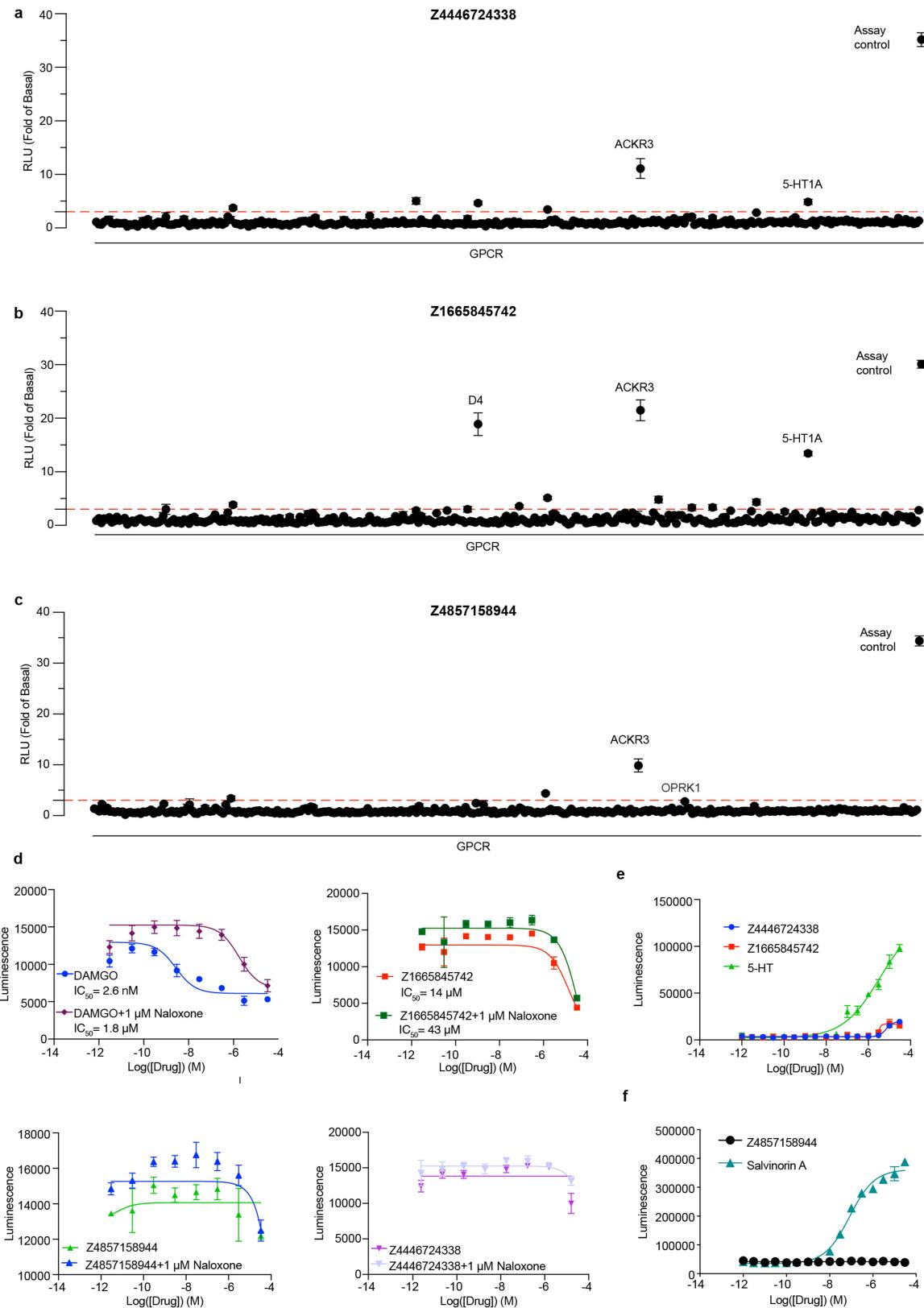
**Extended Data Fig. 4** | See next page for caption.

# Article

## Extended Data Fig. 4 | Effect of systemic $\sigma$ receptor ligands on motor behaviour.

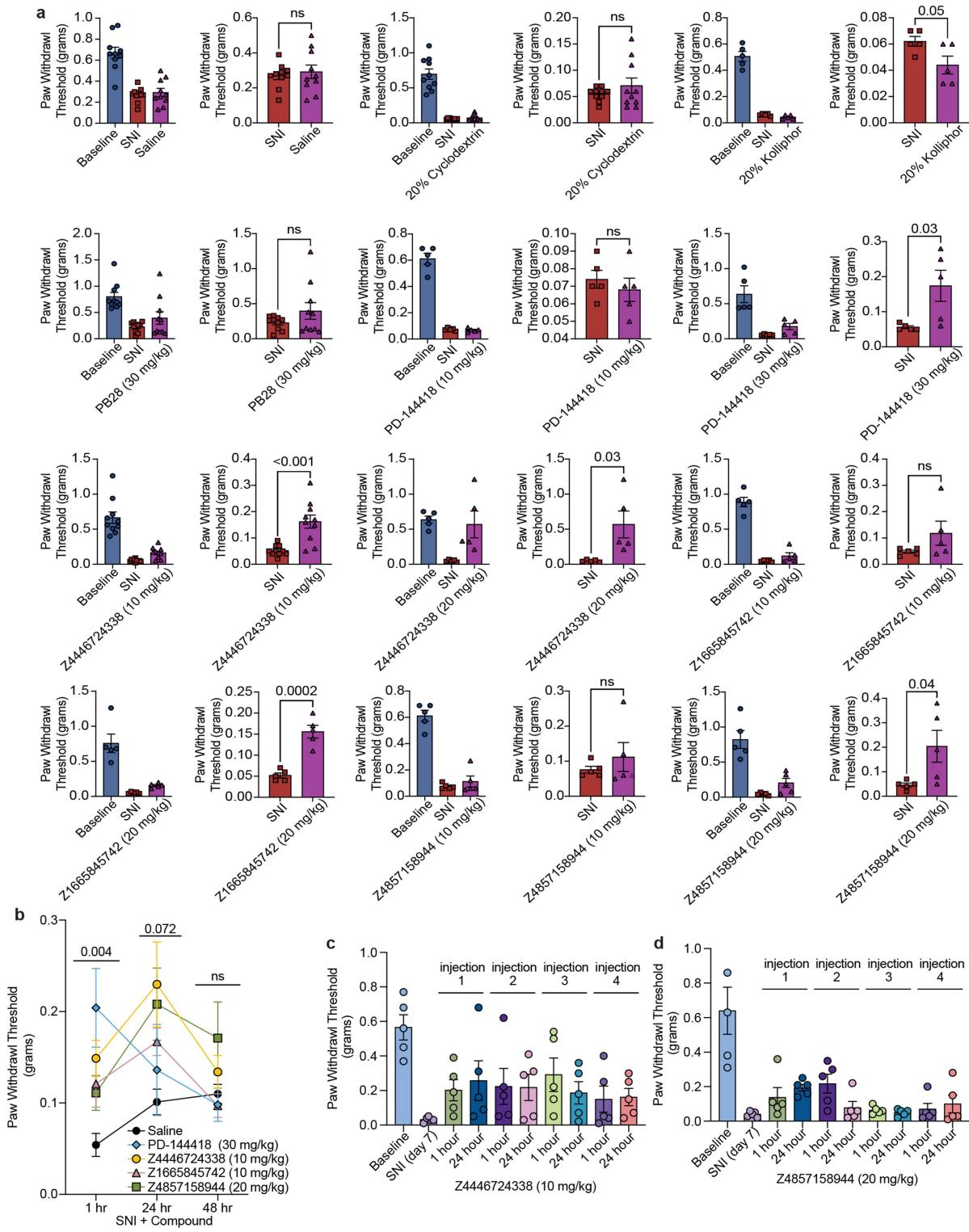
**a.** Response of mice to a von Frey filament after spared nerve injury (SNI). All five ligands are compared to their respective vehicles (PD-144418 10 mg/kg ( $n=5$ ) and 30 mg/kg ( $n=5$ ) vs. kolliphor ( $n=5$ ), one-way ANOVA,  $F(2,12) = 7.49, p = 0.008$ ; Z4446724338 10 mg/kg ( $n=10$ ) and 20 mg/kg ( $n=5$ ) vs cyclodextrin ( $n=10$ ), one-way ANOVA,  $F(2,22) = 25.12, p < 0.001$ ; Z4857158944 10 mg/kg ( $n=5$ ) and 20 mg/kg ( $n=5$ ) vs cyclodextrin ( $n=10$ ), one-way ANOVA,  $F(2,17) = 5.10, p = 0.02$ ; Z1665845742 10 mg/kg ( $n=10$ ) and 20 mg/kg ( $n=5$ ) and PB28 30 mg/kg ( $n=10$ ) vs saline ( $n=10$ ), one-way ANOVA,  $F(3,31) = 6.18, p = 0.002$ ; asterisks define individual group differences to respective vehicle control using Dunnett's multiple comparisons Post-hoc test; ns = not significant,  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ). Data shown are mean  $\pm$  s.e.m. Data for higher doses and vehicles is replotted from Fig. 4. **b.** No sedation or motor impairment on the rotarod was observed after drug treatments compared to vehicle at 1 h (Z1665845742 10 mg/kg ( $n=5$ ) and Z4857158944 20 mg/kg ( $n=5$ ) vs saline ( $n=5$ ), one-way ANOVA,  $F(2,12) = 1.04, p = 0.38$ ; Z4446724338 10 mg/kg ( $n=5$ ) vs kolliphor ( $n=5$ ), unpaired two-tailed Student's  $t$ -test,  $t(8) = 0.47, p = 0.65$ ) or 24 h post-injection (Z1665845742 10 mg/kg ( $n=5$ ) and Z4857158944 20 mg/kg ( $n=5$ ) vs saline ( $n=5$ ), one-way

ANOVA,  $F(2,12) = 0.45, p = 0.65$ ; Z4446724338 10 mg/kg ( $n=5$ ) vs kolliphor ( $n=5$ ), unpaired two-tailed Student's  $t$ -test,  $t(8) = 0.72, p = 0.49$ ); ns = not significant. Data shown are means  $\pm$  s.e.m. **c.** Response of SNI mice to a von Frey filament after repeated injections of Z4446724338 10 mg/kg ( $n=5$ ). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are means  $\pm$  s.e.m. normalized to each mouse's SNI baseline. **d.** Response of SNI mice to a von Frey filament after repeated injections of Z4857158944 10 mg/kg ( $n=5$ ). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are means  $\pm$  s.e.m. normalized to each mouse's SNI baseline. **e.** Quantification of the expression levels of *Sigmar1* ( $\sigma_1$ ) and *Tmem97* ( $\sigma_2$ ) in wildtype (WT) and SNI mice detected by *in situ* hybridization ( $n = 3$  mice per group). Representative images can be found in panel **f**. Data shown are mean  $\pm$  s.e.m.; unpaired two-tailed Student's  $t$ -test—*Sigmar1*:  $t(4) = 0.5, p = 0.64$ ; *Tmem97*:  $t(4) = 1.0, p = 0.37$ ; ns = not significant. AU = arbitrary units. **f.** *in situ* hybridization of mouse dorsal root ganglion (DRG) sections for *Sigmar1* ( $\sigma_1$ ) and *Tmem97* ( $\sigma_2$ ) genes illustrates expression in myelinated (Nefh-positive; blue) and unmyelinated (Acpp-positive; red) subsets of sensory neurons and no change after SNI.



**Extended Data Fig. 5 | Off-target profiling of Z4446724338, Z1665845742 and Z4857158944.** **a–c**, TANGO screens against a panel of 320 GPCRs for the indicated  $\sigma_2$  ligand. **a**, Z4446724338. **b**, Z1665845742. **c**, Z4857158944. **d**, GloSensor  $\mu$ OR-mediated cAMP inhibition ( $G_i$  activation) by DAMGO,

Z4446724338, Z1665845742, and Z4857158944. **e, f**, Follow-up dose-response curves for pain-related receptors that showed activation in **a–c**. **e**, Z4446724338 and Z1665845742 against SHT1A. **f**, Z4857158944 against  $\kappa$ OR. Data shown are means  $\pm$  s.e.m.



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 6 | Paw withdrawal thresholds.** **a**, Paw withdrawal thresholds (PWT) before (blue bar) and after (red bar) SNI, as well as after SNI + treatment (purple bar). For easier visualization of individual data points, data was also plotted without the pre-SNI baseline. Data are the same as in Fig. 4b and Extended Data Fig. 4a, but without the normalization to the individual post-SNI baselines and are expressed as mean  $\pm$  s.e.m.; mice per group: saline ( $n=10$ ); cyclodextrin ( $n=10$ ); kolliphor ( $n=5$ ); PB28 30 mg/kg ( $n=10$ ); PD-144418 10 mg/kg ( $n=5$ ) and 30 mg/kg ( $n=5$ ); Z4446724338 10 mg/kg ( $n=10$ ) and 20 mg/kg ( $n=5$ ); Z1665845742 10 mg/kg ( $n=5$ ) and 20 mg/kg ( $n=5$ ); Z4857158944 10 mg/kg ( $n=5$ ) and 20 mg/kg ( $n=5$ ); unpaired two-tailed Student's t-test. **b**, PWTs 1 h, 24 h, and 48 h after saline or drug treatment. Data are the same as in Fig. 4c, but without the normalization to the individual post-SNI baselines, and are expressed as mean  $\pm$  s.e.m. Significance levels

determined using Dunnett's multiple comparisons Post-hoc test reflect the difference between Z4446724338 and saline for simplicity (two-way ANOVA; time  $\times$  treatment interaction:  $F(8, 80) = 2.4, p = 0.02$ ; time:  $F(2, 74) = 5.2, p = 0.009$ ; treatment:  $F(4, 40) = 3.3, p = 0.02$ ; four treatment groups ( $n=10$ ) except PD-144418 ( $n=5$ ); ns = not significant. **c**, Response of SNI mice to a von Frey filament after repeated injections of Z4446724338 10 mg/kg ( $n=5$ ). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are paw withdrawal thresholds in grams, expressed as mean  $\pm$  s.e.m. **d**, Response of SNI mice to a von Frey filament after repeated injections of Z4857158944 10 mg/kg ( $n=5$ ). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are paw withdrawal thresholds in grams, expressed as mean  $\pm$  s.e.m.

# Article

**Extended Data Table 1 | Data collection and refinement statistics**

	PB28-bound (Se-labeled)	Roluperidone-bound (Native)	Z1241145220-bound (Native)	Z4857158944-bound (Native)	Cholesterol-bound (Native)
<b>Data collection</b>					
Space group	$P2_1$	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_1$
Number of crystals	1	1	1	1	1
Cell dimensions					
a, b, c (Å)	70.6, 55.2, 93.0	69.1, 54.2, 99.7	55.4, 61.5, 110.4	70.7, 55.4, 93.0	70.9, 54.9, 93.0
$\alpha, \beta, \gamma$ (°)	90, 95.0, 90	90, 91.1, 90	90, 90, 90	90, 94.5, 90	90, 94.3, 90
Wavelength (Å)	1.255	1.03320	1.03321	1.033167	1.03320
Resolution (Å)	33.88 - 2.942 (3.047 - 2.942)	42.61 - 2.71 (2.81 - 2.71)	49.5 - 2.41 (2.55 - 2.41)	40.2 - 2.41 (2.55 - 2.41)	47.24 - 2.8 (2.9 - 2.8)
$R_{\text{sym}}$	24.75 (88.16)	26.11 (205.9)	18.4 (177.4)	19.67 (227.6)	32.2 (220.2)
$I/\sigma I$	5.73 (0.93)	5.90 (0.71)	7.50 (0.7)	5.18 (0.56)	5.35 (0.94)
Completeness (%)	98.67 (90.76)	99.54 (99.87)	99.55 (99.36)	97.9 (88.3)	99.7 (98.5)
Redundancy	4.0 (3.4)	6.8 (6.5)	6.2 (4.4)	4.4 (4.5)	6.7 (6.9)
CC <sub>1/2</sub>	98.7 (49.5)	99.4 (36.6)	99.6 (26.9)	99.5 (28.3)	99.2 (42.6)
<b>Refinement</b>					
Resolution (Å)	2.94	2.71	2.41	2.41	2.8
No. reflections	15228	20340	15165	27448	17720
No. reflection used for $R_{\text{free}}$	1524 (10%)	2004 (9.85%)	1063 (7%)	1370 (5%)	1752 (9.89%)
$R_{\text{work}} / R_{\text{free}}$	20.39 / 24.26	22.18 / 25.2	21.36 / 24.6	25.0 / 28.8	24.06 / 27.81
No. atoms					
Protein	5490	5472	2761	5393	5292
Lipid/ion	231	250	148	231	223
Ligand	108	108	48	100	112
Water	27	7	46	37	21
B-factors (Å <sup>2</sup> )					
Protein	49.68	67.89	50.19	57.24	64.29
Lipid/ion	52.48	65.57	59.49	62.32	62.33
Ligand	56.89	79.39	49.32	66.07	71.96
Water	45.66	62.72	57.08	57.77	61.02
R.m.s. deviations					
Bond lengths (Å)	0.003	0.003	0.005	0.003	0.003
Bond angles (°)	0.61	0.61	1.04	0.58	0.59

**Extended Data Table 2 | 14 of the highest-affinity direct docking hits for the  $\sigma_2$  receptor**

2D drawing	ZINC ID	Rank	DOCK score (kcal/mol)	TC*	$K_i$ (nM)		Selectivity ( $\sigma_1/\sigma_2$ )
					$\frac{K_i}{\sigma_2}$	$\frac{K_i}{\sigma_1}$	
	ZINC000450573233	4429	-57.25	0.32	4.3	128	30
	ZINC000895657866	19047	-55.35	0.31	21.4	989.6	46
	ZINC001170548029	4945	-57.11	0.35	22.6	727.2	32
	ZINC000533478938	18545	-55.38	0.30	34.5	1470	43
	ZINC000921927365	983	-59.01	0.31	67.3	1186	18
	ZINC000548355486	7007	-56.68	0.29	2.4	4.9	2
	ZINC000348332392	931	-59.07	0.28	33.7	2.9	0.1
	ZINC001254761628	16059	-55.58	0.27	4.7	53	11.3
	ZINC000544117725	3522	-57.52	0.28	10	16.25	1.6
	ZINC000170908795	13281	-55.84	0.29	6.7	32.7	4.9
	ZINC001196519317	9290	-56.3	0.29	2.4	13.4	5.6
	ZINC000656714762	1276	-58.68	0.26	67.8	4.6	0.1
	ZINC001237901728	11409	-56.03	0.30	27	1.6	0.1
	ZINC001460312963	11817	-55.99	0.29	5.2	1.7	0.3

See Supplementary Table 1 for all 484 compounds tested.

\*TC, Tanimoto coefficient to sigma ligands from ChEMBL.

# Article

**Extended Data Table 3 | Measured pharmacokinetic parameters for PB28, Z1665845742, Z4446724338 and Z4857158944 in male CD-1 mice by 10 mg kg<sup>-1</sup> subcutaneous administration**

Pharmacokinetic Parameters							
Type	Name	T <sub>max</sub> min	C <sub>max</sub> ng/ml (g)	AUC <sub>0→t</sub> (AUC <sub>last</sub> ) ng*min/ml (g)	AUC <sub>0→∞</sub> (AUC <sub>inf_obs</sub> ) ng*min/ml (g)	T <sub>1/2</sub> min	K <sub>el</sub> min <sup>-1</sup>
Plasma	Z1665845742	20	968	99000	112000	185	0.00374
	Z4446724338	20	449	58300	60500	47.4	0.0146
	Z4857158944	20	228	13300	14200	27.8	0.0249
	PB28	60	42	8640	45900	740	0.000937
Brain	Z1665845742	20	3150	436000	509000	747	0.000928
	Z4446724338	20	7390	1140000	1150000	69.7	0.00995
	Z4857158944	20	2960	247000	327000	452	0.00153
	PB28	60	948	229000	240000	98.1	0.00706

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	Data collection was performed at APS beamlines 23-ID-B and 23-ID-D. All programs typically used at these beamlines were used. Docking was done using DOCK3.7.2
Data analysis	XDS (version 31 Jan) HKL2000 v721.2. Coot version 0.9.5 Phenix 1.19.2-4158 PyMOL v2.5 UCSF Chimera 1.15 GraphPad Prism 9 R Macromodel (2019 release)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The structures presented in this study have been deposited in the Protein Data Bank (PDB) under the PDB accession codes 7M93, 7M94, 7M95, 7M96, 7MFI. The compounds docked in this study are freely available from our ZINC database, <http://zinc15.docking.org>. All active compounds are available either from the authors or may be purchased from Enamine.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We did not perform sample-size calculations. We modeled our sample sizes for behavioral studies on previous studies using a similar approach to our own, which have been demonstrated to be capable of detecting significant changes (Scherrer et al, 2009; Muralidharan, A. et al 2021).
Data exclusions	No data were excluded.
Replication	All replications were successful at least twice.
Randomization	The animals were randomly assigned to the treatment group and control group. For behavioral experiments, animals were initially placed into one cage and allowed to free run for a few minutes. Next, each animal was randomly picked up, injected with the drug or vehicle control and placed into a separate cylinder before the behavior test. All experiments were for animal behavior and followed this randomization protocol.
Blinding	For all behavioral testing the experimenter was always blind to treatment. All experiments were in animals and under blinding conditions.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Expi293 cells were from Thermo Fisher Scientific. sf9 cells were from Expression Systems. HEK-293T cells were from ATCC.

Authentication

Cell lines were not authenticated. All cells used in this study are commercial and were obtained from vendors as indicated in the manuscript.

Mycoplasma contamination

Cells were confirmed to be mycoplasma free

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Behavioral testing was performed on adult (8-10 weeks old) male C56BL/6 mice (strain #664) purchased from the Jackson Laboratory. The housing conditions used a 12 hour light on-light off schedule. Temperature 71.6oF; humidity 58%.

Wild animals

No wild animals were used in this study.

Field-collected samples

No field collected samples were used in the study.

Ethics oversight

All animal experiments were approved by the Institutional Animal Care and Use Committee at UCSF and were conducted in accordance with the NIH Guide for the Care and Use of Laboratory animals

Note that full information on the approval of the study protocol must also be provided in the manuscript.