

Development of a random background to understand ligand optimization

<https://doi.org/10.1038/s41586-026-11013-5>

Received: 5 July 2025

Accepted: 10 August 2026

Published online: 16 September 2026

Open access

 Check for updates

Xinyu Xu^{1,14}, Olivier Mailhot^{1,14}, Galen J. Correy^{2,14}, Xi-Ping Huang^{3,14}, Joao M. Braz⁴, Da Shi⁵, Karthik Srinivasan¹, Kara Zielinski², Yuliia Holota⁶, Yuliia Kuziv⁶, Christos Iliopoulos-Tsoutsouvas⁷, Nathan D. Levinzon⁸, Yagmur U. Doruk⁹, Moira M. Rachman¹, Morgan E. Diolaiti⁹, Maisie G. V. Stevens⁹, Fangyu Liu¹, Katie L. Holland¹, Harald Hübner¹⁰, Jing Wang³, Yujin Wu¹, Alan Ashworth⁹, Alexandros Makriyannis⁷, Yuqi Zhang⁵, Yurii S. Moroz^{6,11,12}, Peter Gmeiner¹⁰, Robert Abel¹³, Aashish Manglik¹, Allan I. Basbaum⁴, Bryan L. Roth^{3,14}, James S. Fraser^{2,14} & Brian K. Shoichet^{1,14}

Ligand optimization is central to drug discovery, with hundreds of analogues often designed and synthesized between an initial hit and a therapeutic candidate^{1,2}. The efficiency of this process is unclear, partly because there is no random background for optimization to compare against. Such a random background might emerge from systematic random small substitutions across starting ligands, measuring the likelihood of achieving a substantial improvement in affinity or potency, or other property by any single perturbation. Recent literature has suggested that perhaps 10% of analogues with minor modifications improve upon the potency of a parent by tenfold or more^{3,4}, but this number is clouded by reporting bias, intentional improvement and inter-group variability. To begin to establish a background expectation for ligand optimization, here we systematically modified 18 lead molecules across six targets with single-atom changes; 257 compounds were synthesized. Unexpectedly, 11.3% of these random small perturbation analogues improved potency by tenfold or more. Conversely, they typically had worse in vitro pharmacokinetics. Although it was possible to find analogues where the potency increase compensated for inferior exposure and half-life, resulting in more potent compounds in vivo, overall, a frustrated landscape for ligand optimization is revealed. This study begins to establish a background expectation for ligand potency optimization and offers a simple strategy to do so. It also begins to quantify the challenges confronting the field in moving beyond in vitro potency.

Ligand optimization is central to chemical probe and drug discovery, as initial active molecules rarely have the potency or pharmacokinetic (PK) properties to be viable in vivo^{1,2}. Between the discovery of an initial hit and a clinical candidate many hundreds, occasionally thousands of optimized analogues might be synthesized^{1,5}. Several strategies^{6–10} have been developed to improve this process, ranging from early empirical approaches such as Topliss Trees¹¹ and quantitative structure–activity relationships¹² to contemporary artificial intelligence (AI)-guided PK¹³ prediction and free energy calculations^{14,15}. The efficiency of these strategies remains uncertain as there is no random background against which to compare them. Using tenfold improvement between analogue and parent as a benchmark for substantial impact, 10% of analogues meet this standard in the ChEMBL database¹⁶ (Extended Data Fig. 1a). As these ChEMBL results suffer from success bias, sample multiple types

of perturbation and suffer from inter-group irreproducibility¹⁷, the public domain offers no sure guidance on what level of improvement one might expect in ligand optimization from random conservative changes.

Random backgrounds have long been used in biology to quantify significance. In genomics, they distinguish between artificial and natural selection¹⁸. In epidemiology, random incidence distinguishes genetic diseases from those driven by environmental factors^{19,20}. In protein engineering, random backgrounds help to evaluate improvements in enzyme function²¹, and alanine scanning introduces a minimal perturbation to find hot spots for binding and function²². A random background for ligand optimization might quantify progress across chemotypes and targets and compare different optimization strategies.

¹Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA, USA. ²Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA. ³Department of Pharmacology, NIMH Psychoactive Drug Screening Program, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁴Department of Anatomy, University of California San Francisco, San Francisco, CA, USA. ⁵Schrödinger Inc., San Diego, CA, USA. ⁶Enamine Ltd., Kyiv, Ukraine. ⁷Center for Drug Discovery, Department of Pharmaceutical Sciences, Northeastern University, Boston, MA, USA. ⁸Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, UT, USA. ⁹Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA. ¹⁰Department of Chemistry and Pharmacy, Medicinal Chemistry, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. ¹¹Chemspace LLC, Kyiv, Ukraine. ¹²Taras Shevchenko National University of Kyiv, Kyiv, Ukraine. ¹³Schrödinger Inc., New York, NY, USA. ¹⁴These authors contributed equally: Xinyu Xu, Olivier Mailhot, Galen J. Correy, Xi-Ping Huang. ✉e-mail: bryan_roth@med.unc.edu; jaimefraser@gmail.com; bshoichet@gmail.com

Table 1 | Analogues from all parents improving affinity or potency by threefold or more or tenfold or more

Protein	Assay readout	Parent ligand	Parent affinity or potency	Total analogues	Analogues with threefold or more affinity or potency improvement	Analogues with tenfold or more affinity or potency improvement
μOR	EC ₅₀	Z4407498716	9.2 nM	16	2	2
α ₂ A	EC ₅₀	Z2750653629	1.4 nM	31	7	4 ^a
α ₂ A	K _i	Z3034773248	22.1 μM	18	3	1
α ₂ A	EC ₅₀	Z4376630014	8.3 nM	10	2	0
α ₂ A	K _i	Z8727395870	13 μM	18	1	0
AmpC	K _i	Z2610488449	24.2 μM	10	2	0
CB2	K _i	Z52076138	2.3 μM	7	7	6 ^b
CB2	K _i	Z6969215903	0.4 μM	6	0	0
CB2	EC ₅₀	Z8184698918	0.1 μM	20	2	2
Mac1	IC ₅₀	Z5398393122	2.2 μM	5	0	0
Mac1	IC ₅₀	Z1039063794	0.7 μM	12	0	0
Mac1	IC ₅₀	Z7534253453	8.9 μM	22	8	2
SERT	K _i	Z2009978218	1.1 μM	14	4	0
SERT	K _i	Z2573292480	0.08 μM	11	1	0
SERT	K _i	Z2573292509	0.2 μM	9	2	1
SERT	K _i	Z6971277399	13.4 μM	15	10	4
SERT	K _i	Z8727393896	42.7 μM	14	8	3 ^c
SERT	K _i	Z8731642686	1.8 μM	19	10	4
Total		All parents		257	69	29

Fold changes are displayed as nearest-integer values; exact (unrounded) fold changes are provided in Supplementary Table 1. Using exact fold changes, 62 analogues improve threefold or more and 25 analogues improve tenfold or more. ^aOne of these had a fold change of 9.7. ^bTwo of these had fold changes of 9.7 and 9.8. ^cOne of these had a fold change of 9.9 (see Supplementary Table 1).

Like alanine scanning, an unbiased background for ligand optimization might involve conservative perturbations, systematically and comprehensively applied across multiple parents with different targets. This resembles a ‘positional analogue scanning’ approach that substitutes ligand non-polar hydrogens with groups such as CH₃, OH, Cl, F and Br, and aromatic carbon atoms with nitrogen²³. In a retrospective analysis of 110,000 matched molecular pairs in the ChEMBL24 database^{3,4}, about 30% of analogues improved over threefold in affinity or potency versus their parent. This observation is tempered by design and reporting biases towards compound improvement, by the variability in experimental conditions across datasets that can reduce reproducibility¹⁷, and by the few parent compounds that had multiple conservative changes. Thus, it seemed interesting to create a set of parent ligands that were systematically and comprehensively modified by small perturbations, without design and without a bias towards improvement, using the positional analogue scanning approach^{23,24} where the differences between parent and the suite of analogues were quantified by a single laboratory.

To create what we refer to as a random background set for ligand optimization, we chose 18 parent ligands spanning six targets, including three GPCRs (α₂A adrenergic receptor (α₂AAR), μ-opioid receptor (μOR) and cannabinoid receptor type 2 (CB2)), one transporter (the serotonin transporter (SERT)) and two soluble enzymes (AmpC β-lactamase (AmpC) and macrodomain 1 of SARS-CoV-2 (Mac1)). This set was chosen for targets that we had under experimental control and is admittedly far from comprehensive. The 18 parents were chosen for similar pragmatic reasons; they were molecules to which we had ready access. Still, these proteins provide a range of target classes while the 18 parents cover a wide range of physicochemical properties (Extended Data Fig. 2): ranging from 1 nM to 43 μM activity (Table 1) and representing 11 unrelated scaffolds. For each parent, we systematically explored all analogues accessible by a single-atom or methyl modification where such modifications were synthetically feasible and where they did not modify the net charge of the parent at physiological pH. The single-atom modifications involved substituting carbon

hydrogens with one of methyl²⁵, hydroxyl²⁶, chloro²⁷, fluoro²⁸ or rarely bromo, or substituting aromatic carbon atoms with nitrogens (Fig. 1). These modifications are not comprehensive but do represent those preferred by medicinal chemists.

These substitutions were systematically and comprehensively applied without design, ensuring an approach unbiased by anything other than synthetic accessibility and price. Overall, 257 single-atom analogues were synthesized and tested for changes in activity on the target, allowing us to measure how often we might expect affinity or potency to improve tenfold or more by minimal perturbation. This begins to provide a background expectation for how often a designed analogue might meaningfully improve affinity or potency. As we show, these undesigned changes had an unexpectedly high success rate, suggesting a systematic strategy for optimization, even though that was not our goal. As affinity or potency maturation is only one criterion for advancing from hit to candidate, we also measured *in vitro* PK properties for the analogues, including permeability, metabolic stability, plasma protein binding and three other terms, versus the parent compounds. These studies explore trade-offs between PK and target affinity or potency in ligand optimization.

Systematic small-change analogues

We began with 18 parent ligands, most of which are unrelated to one another topologically—that is, by their patterns of atoms and bonds (ECFP4 Tanimoto coefficients ≤ 0.35), with affinities ranging from 1 nM to 43 μM and molecular weights ranging from 200 to 400 amu (Extended Data Fig. 2). Each of the 18 parents was modified one atom at a time with five types of single-atom substitutions (Fig. 1a, central panel), subject only to expense constraints (all molecules cost US\$400 or less). About 7% of the possible molecules met these criteria and of those approximately 80% were successfully synthesized, resulting in 257 analogues. In the analyses that follow, we pooled all 257 analogues for analysis, improving statistical power (Supplementary Fig. 1).

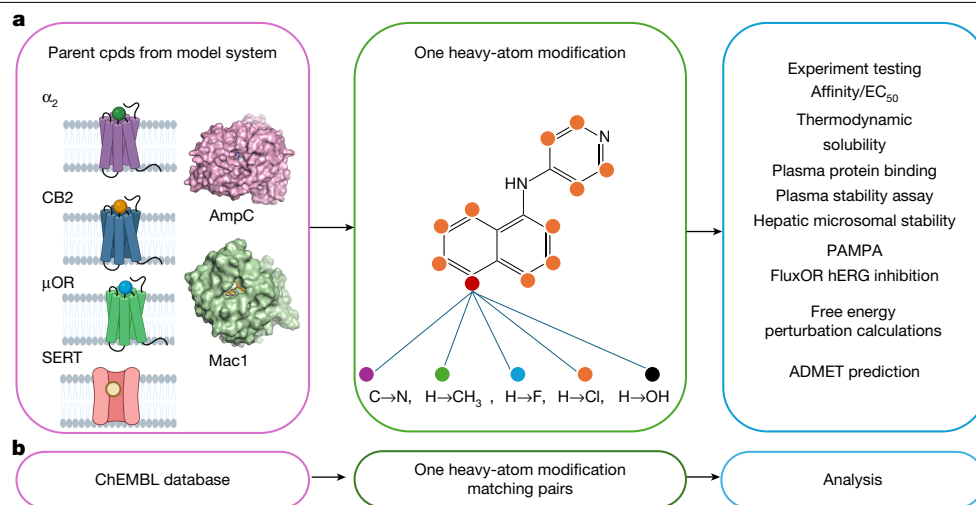


Fig. 1 | Experimental and analysis workflow for one heavy-atom analogues.

a, Parent ligands from six drug targets (the aminergic GPCR α_2 AAR, the lipid GPCR CB2R, the peptide GPCR μ OR, the serotonin SLC transporter SERT, and the soluble enzymes AmpC and Mac1) were modified by one-at-a-time, one-atom substitutions: H→CH₃, H→Cl, H→OH, H→F and ring C→N. These modifications were made systematically and comprehensively; only compounds that were too expensive to synthesize or that changed the charge state of the parent were left out. The analogues were tested for changes in

affinity or potency, solubility, plasma protein binding, plasma stability, hepatic microsomal stability, permeability and hERG inhibition, versus their parent ligands. Relative affinity and in vitro PK changes were also evaluated computationally using FEP+ and ADMET prediction models. **cpd**, compound. **b**, For comparison with the experimental set, matched parent-analogue pairs differing by one heavy-atom modification were identified in the ChEMBL database and analysed using the same activity-improvement framework. Figure created in BioRender; Xu, X. <https://biorender.com/b6zqsv7> (2026).

Affinity or potency changes between parents and the small-change analogues

The 257 analogues were tested for activity changes (dissociation constant (K_d), half-maximal inhibitory concentration (IC₅₀) or half-maximal effective concentration (EC₅₀)) versus the parents across the six targets. For μ OR, EC₅₀ values were determined in cAMP Glo-sensor assays. α_2 A and CB2 ligands were assessed by cAMP assays and radioligand competition binding, while SERT activity was assessed using a transporter uptake assay. AmpC inhibition was measured enzymatically and that of Mac1 was measured by ligand displacement. For every parent, we have only reported changes for one type of activity. For instance, if a parent inhibitory constant (K_i) is reported, then only K_i values are reported for its analogues, and if a parent is an agonist, only EC₅₀ values are reported for its analogues. Each functional or inhibition assay included a literature positive-control molecule (Supplementary Fig. 2) with representative concentration–response curves and Z values for each assay (Supplementary Fig. 3). The effects of receptor expression on agonism were controlled for (Supplementary Figs. 4 and 5). Full concentration–response curves and analyses are shown for every parent-analogue pair, allowing for direct inspection.

Overall, 23% of the analogues affected activity by twofold or less (fold changes rounded to the nearest integer; Supplementary Table 1) and another 30% reduced activity by tenfold or more (Extended Data Fig. 1b). Intriguingly, 29 of the 257 analogues (11.3%) improved activity by tenfold or more (Table 1). We note that four of these 29 had fold changes of between 9.7 and 9.9; we consider these effectively tenfold improvements and used the same convention in analysing the public ChEMBL data (see Table 1 and Supplementary Table 1 for details). Such tenfold or more improvement analogues were found for five of the six targets, and for 10 of the 18 parents (Table 1). The effects and chemical identities of all analogues versus their parents are listed in Supplementary Table 1.

We were interested in which perturbations had the biggest effects, how the rates of affinity or potency improvements compared with the literature, with its admitted biases, and if physical properties of the parents or of the binding sites correlated with greater likelihoods of improvement. Among the analogues, methyl substitutions were the

most likely to improve activity tenfold or more, with chlorine substitutions the second most likely; for threefold or more improvements, of which there were 69 among the 257 analogues, both chlorine and methyl substitutions were the most common (Fig. 2a). Given the similar size and similar hydrophobicity of methyls and chlorines, their similar effects seem sensible. Conversely, fluorine substitutions, which are often used to increase metabolic stability, yielded only two tenfold or more and five threefold or more activity improvements after rounding fold changes to the nearest integer. Aromatic carbon-to-nitrogen substitutions also rarely improved affinity or potency, although they often had the greatest effects on PK properties.

Because hydrophobicity increases will increase affinity simply by disfavoring the unbound state, and will have physical property liabilities, it is useful to consider the improved affinity or potency of the analogues in light of ligand efficiency and lipophilic efficiency (LipE). If affinity or potency improvements were largely driven by hydrophobicity, we would expect ligand efficiency to track with hydrophobicity (here, the calculated octanol–water partition coefficient (clog *P*)) and LipE to remain constant or deteriorate. Instead, analogues that improved tenfold or more had ligand efficiencies that increased much more than hydrophobicity, and had substantially increased LipEs. For instance, an analogue with an added methyl, chloro or other atom that improved in binding or in EC₅₀ by even threefold had a ligand efficiency of 0.65 kcal per atom for the added atom, well above the 0.3 kcal per atom thought to be useful for drug optimization^{29,30}, whereas for analogues that improved activity tenfold or more, the single-atom ligand efficiency was 1.4 or more, close to the efficiency limit³¹. Similarly, every analogue that improved activity by more than tenfold also improved LipE (Extended Data Fig. 3f). More broadly, among the 220 analogue–parent pairs with complete numerical activity data, Δ clog *P* showed little quantitative correlation with Δ pActivity (Extended Data Fig. 3a).

We investigated whether there were correlations between likelihood of potency improvement and ligand and binding-site properties. For instance, one might expect that weaker ligands might be easier to optimize, or that smaller binding sites might be better suited to big affinity jumps. We calculated whether the parent molecular weight, binding-site volume, the ligandability of the binding sites (calculated using *D*-score³²) or the affinity or potency of the starting parent

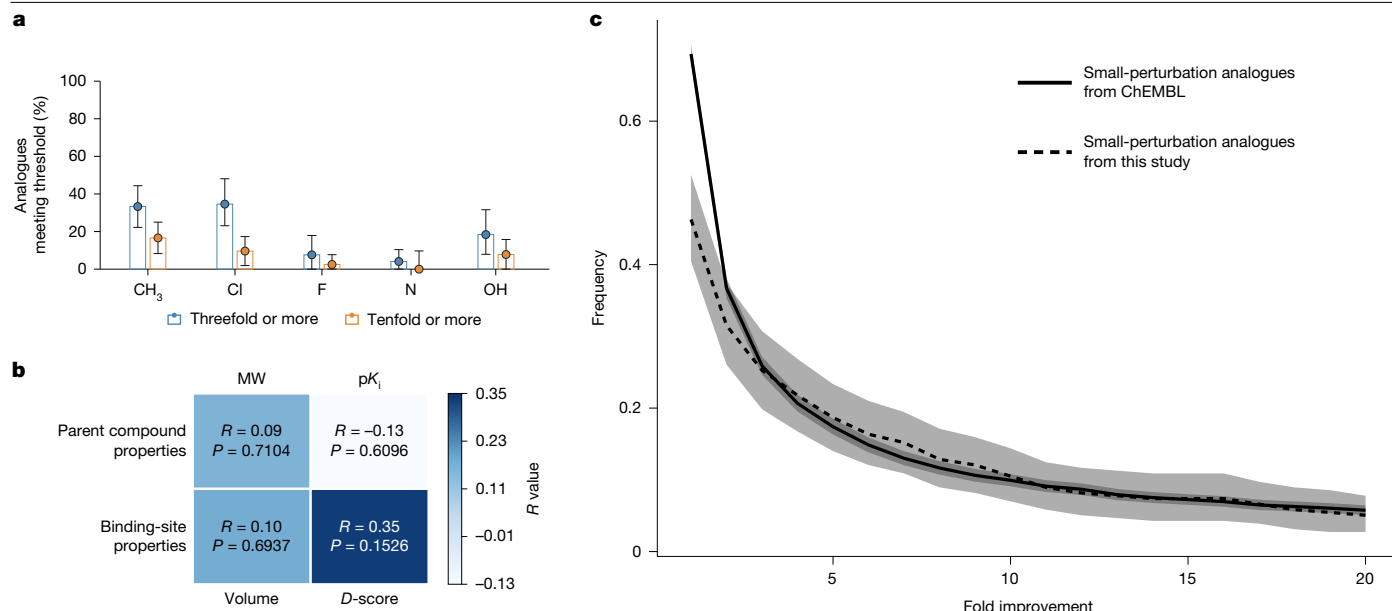


Fig. 2 | Property correlations of affinity or potency improvement and comparison with literature values. **a**, Observed percentages of compounds achieving threefold or more or tenfold or more activity improvement by single modification (CH₃, Cl, F, N and OH). The open bars and points show observed percentages; the black error bars show two-sided 95% confidence intervals estimated by percentile bootstrap (10,000 resamples). Sample sizes were $n = 72, 52, 39, 48$ and 38 analogues for CH₃, Cl, F, N and OH, respectively. **b**, Pearson product-moment correlations between the parent-level rate of analogues achieving tenfold or more affinity or potency improvement and parent molecular weight (MW), parent pK_i , binding-site volume and D -score

($n = 18$ parent compounds for each correlation). R denotes Pearson's correlation coefficient. P values are two-sided: $P = 0.7104, 0.6096, 0.6937$ and 0.1526 for MW, pK_i , binding-site volume and D -score, respectively. **c**, The fraction of single-atom modifications (CH₃, F, Cl, N and OH) that improve activity, comparing the experimental observations in this study with results from the ChEMBL database. The lines show the observed frequencies; the shaded regions indicate two-sided 95% confidence intervals estimated from 10,000 iterations of bootstrap resampling. The ChEMBL ligands are a composite of 54,021 parents and 79,984 small-change analogues; for only 4.7% of these were there more than three analogues per parent.

compound was correlated with the likelihood of finding analogues with substantial affinity or potency improvement. Using the fraction of analogues achieving tenfold or more improvement as the outcome, none of these properties was meaningfully correlated (Fig. 2b). If one considers the magnitude of the affinity or potency change overall, and not just whether it improved tenfold or more, correlations did emerge, with weaker versus stronger parents more likely to support improved affinity ($R = -0.34, P = 3.4 \times 10^{-7}$) and with larger binding sites also tending to do so ($R = 0.24, P = 4.0 \times 10^{-4}$; Supplementary Fig. 6). Although these trends were echoed among matched pairs in ChEMBL (Fig. 2c and Extended Data Fig. 1), the correlations overall were weak. Together, these results suggest that: (1) as a strategy, systematic small perturbations can be unexpectedly successful at improving ligand affinity or potency^{23,24}, perhaps even compared with design-heavy approaches³³; (2) this effect is common across diverse parents and binding sites; (3) it can find echoes in the literature; and (4) random expectations for substantial affinity or potency improvement by small perturbation may be as high as 11% (95% CI 7.5–15.2%).

Structural bases of activity changes

To investigate the structural bases for the affinity or potency changes, we determined the structure of the parent inhibitors in complex with Mac1 and six analogues from the '3122 series, 12 from the '3453 series and four from the '3794 series, each to between 0.97 and 1.02 Å resolution (Extended Data Tables 1–4). The analogues with structures determined represented all six types of single-atom modifications in this study; 15 of the analogues differed in affinity or potency from their parents by threefold or more, two by tenfold or more and among themselves by up to 680-fold.

For some parents, analogues superposed well, facilitating analysis. For the Mac1 '3453 (IC₅₀ of 8.9 μ M) inhibitor series (Extended Data

Fig. 4a), the introduced chloro at the ortho-position of the phenyl ring ('0676, IC₅₀ of 21 μ M) projects into solvent and loses activity versus the parent (Extended Data Fig. 4b). Moving this chloro to the meta-position ('1304), where it packs with Phe132, Ile131 and Gly48, improves inhibition 11-fold ($P < 0.01$) to 1.9 μ M (Extended Data Fig. 4c). Moving it one atom further over, to the para-position, had little effect ('6343, IC₅₀ of 2.1 μ M), although a methyl at the same position ('9870, IC₅₀ of 7.6 μ M) loses fourfold affinity ($P < 0.01$), presumably reflecting its poorer packing (Extended Data Fig. 4d).

Many of the effects of the small perturbation analogues could be explained post hoc from their structures; however, fewer were easily anticipated. Even the apparent simplicity of the '3453 series is belied by structural accommodations in the enzyme. For instance, the para-chloro of the 2 μ M '6343 would have clashed with the enzyme without Phe132 and Ile131 rotating away. In doing so, Phe132 adopts a partially eclipsed conformation. Presumably, the resulting strain is overcome by improved ligand packing, but this balance of terms is non-obvious (encouragingly, free energy calculations did correctly predict that the para-chloro and meta-chloro would improve affinity, whereas the ortho-chloro would lose it, although the agreement was more qualitative than quantitative; see below). Conversely, the methyl analogue ('9249, 53 μ M) of '0676 (IC₅₀ of 22 μ M) is buried from solvent and might be expected to be more potent than its chloro cousin, but instead it loses another 2.5-fold in K_i ($P < 0.01$; here, the opposite effect was anticipated by the free energy calculation). Meanwhile, in the complex between '4905 and the α_2 AAR, an added methyl group fits into a pre-existing sub-pocket and qualitatively its improved potency is reasonable (Fig. 5a). Quantitatively, its 52-fold effect on EC₅₀ is at the outer edge of what might be expected, especially as the methyl buries a serine hydroxyl. This is echoed by the free energy calculation, which here too correctly predicted improved activity but by twofold not 52-fold. These are examples of complexes where the analogues

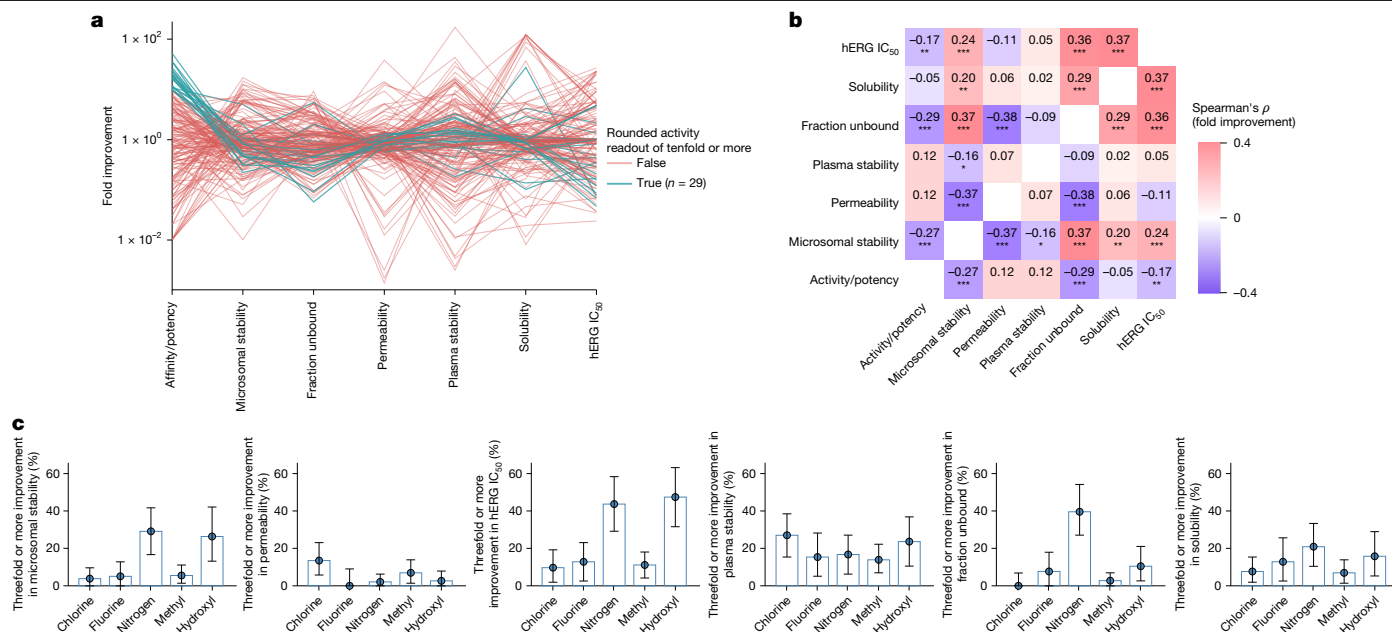


Fig. 3 | A frustrated landscape for ligand optimization. a, Fold changes in PK properties for the 29 analogues classified as tenfold or more improved after integer rounding in affinity or potency versus a parent. **b**, Correlations between affinity or potency fold changes and those of the other PK properties. Correlations were assessed by two-sided Spearman rank correlation; exact pairwise-complete sample sizes (analogue–parent pairs) were: activity or potency versus microsomal stability, permeability, plasma stability, fraction unbound, solubility and hERG IC₅₀, *n* = 203, 203, 203, 193, 191 and 143, respectively; microsomal stability versus permeability, plasma stability, fraction unbound, solubility and hERG IC₅₀, *n* = 240, 240, 213, 220 and 166, respectively; permeability versus plasma stability, fraction unbound, solubility

and hERG IC₅₀, *n* = 240, 213, 220 and 166, respectively; plasma stability versus fraction unbound, solubility and hERG IC₅₀, *n* = 213, 220 and 166, respectively; fraction unbound versus solubility and hERG IC₅₀, *n* = 200 and 143, respectively; and solubility versus hERG IC₅₀, *n* = 154. **P* < 0.05, ***P* < 0.01 and ****P* < 0.001. Exact two-sided *P* values for all correlations are provided in the Source Data. **c**, Observed percentages of single-atom changes achieving threefold or more improvement in each parameter. The open bars and points show observed percentages; the black error bars show two-sided 95% confidence intervals estimated by percentile bootstrap resampling (10,000 resamples). Sample sizes by modification were *n* = 52, 39, 48, 72 and 38 analogues for chlorine, fluorine, nitrogen, methyl and hydroxyl, respectively.

superpose on the parent and each other. In complexes in which the single-atom changes led to substantial inhibitor movement or to the adoption of multiple ligand conformations, prediction was harder still (Extended Data Fig. 4f,g). For example, the Mac1 methyl analogue '3176 repositions the inhibitors in the pocket (Extended Data Fig. 4g), whereas compounds '9249 and '3194 bind in two distinct conformations (Extended Data Fig. 4g), a phenomenon that is probably underreported in the Protein Data Bank³⁴ but is readily seen at the ultra-high resolution of these structures. Together, this set of perturbations may provide an interesting and sometimes challenging test for prediction methods.

A frustrated landscape for ligand affinity or potency and PK

In optimizing chemical probes and leads, compound PK are as important as molecular target affinity or potency^{35,36}. We thus explored how the small-perturbation analogues affected PK properties versus those of the parent ligands, and how these changes related to alterations of affinity or potency. We measured the following in vitro PK properties: metabolic stability in microsomes (half-life), plasma stability (half-life), plasma protein binding (fraction unbound), solubility (μM), hERG inhibition (IC₅₀, μM) and membrane permeability (parallel artificial membrane permeability assay (PAMPA); cm s⁻¹), for both the parent ligands and their small-change analogues (Fig. 3 and Supplementary Tables 2–7). These admittedly in vitro measurements are considered predictive of in vivo PK^{37–39}. Some of the small-change analogues showed meaningful improvements relative to the parent ligands for each of these properties (Supplementary Table 8). For example, 12.4% of the analogues had threefold or more improvement in microsomal stability, 20.8% improved threefold or more in plasma stability, 13.8% improved

solubility by threefold or more, 6.2% improved threefold or more in permeability, and 13.2% of the analogues improved threefold or more in plasma fraction unbound. Thus, each property can be meaningfully improved by simple changes. Correspondingly, a substantial fraction of the analogues experienced threefold or more decreases in these same properties (Extended Data Fig. 5). Considering what are perhaps the three most impactful in vitro PK measurements, microsomal stability, fraction unbound and permeability, 41.4% of the analogues experienced more than threefold losses in at least one of these, almost all (93.1%) experienced at least some deterioration by the same criterion, and 34.5% deteriorated in all three in vitro PK properties.

Although many of the small-perturbation analogues substantially improved individual PK properties, none of the 29 analogues classified as tenfold or more improved after integer rounding showed improvements across all three key properties: metabolic stability, PAMPA permeability and plasma fraction unbound (Fig. 3a and Supplementary Fig. 7); many were worse in all three properties. Indeed, the changes in in vitro PK were, at best, orthogonal to affinity or potency fold change (Fig. 3b), and most were, if anything, anti-correlated with it, including plasma fraction unbound, microsomal stability, hERG inhibition and, to a minor extent, solubility. This is borne out at the level of individual substitutions: the substituent groups that most often improved affinity or potency substantially, such as methyl and chloro groups (Fig. 2a), were least associated with improved PK properties (Fig. 3c and Supplementary Fig. 8). Meanwhile, those changes least likely to improve affinity or potency substantially, such as aromatic C→N, were most likely to improve microsomal stability or fraction unbound (Supplementary Fig. 8). More broadly, most of the PK properties were largely independent of one another, with no correlations exceeding $|\rho| > 0.5$ and those that were statistically significant having Spearman ρ values ranging

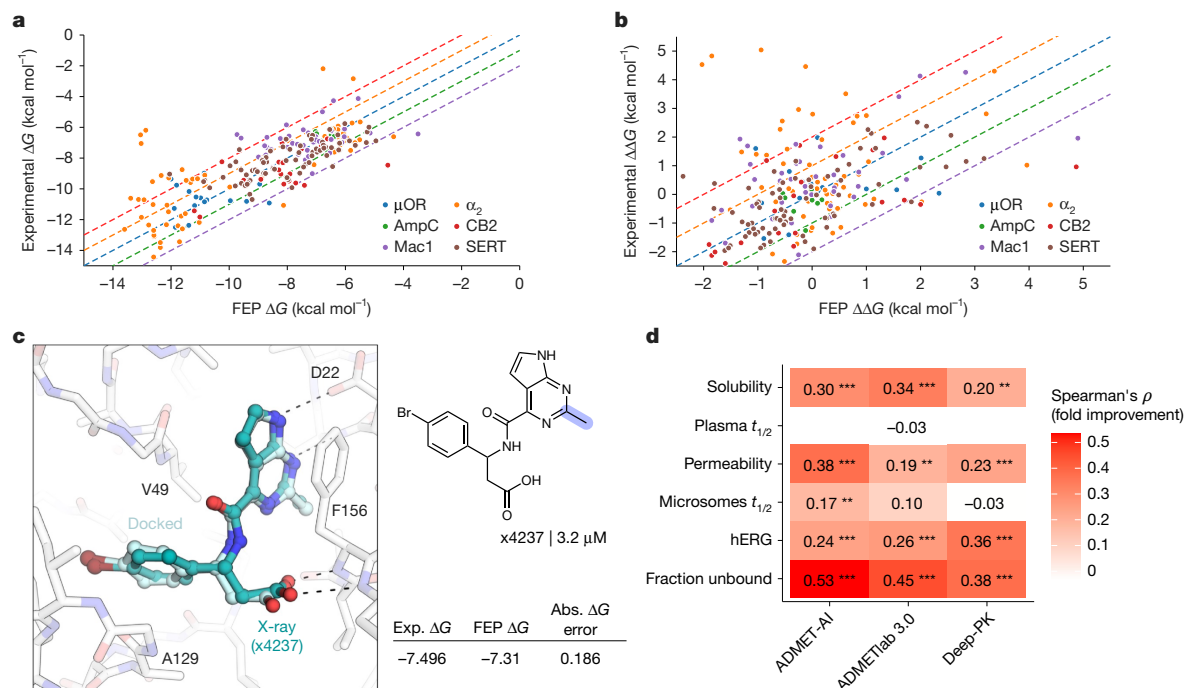


Fig. 4 | Accuracy of computational predictions of analogue relative affinity and in vitro PK. a, Comparison of FEP-predicted and experimentally measured binding free energies (ΔG). The dashed lines indicate ± 1 kcal mol⁻¹ and ± 2 kcal mol⁻¹ error margins. **b**, Correlation between FEP-predicted and experimentally measured relative binding free energy changes ($\Delta\Delta G$). Dashed lines are as in panel **a**. **c**, An example of good structural alignment between the predicted (cyan) and experimental (light blue) poses, accompanied by a small error in the predicted $\Delta\Delta G$. Dashed lines indicate hydrogen-bonding interactions. Abs., absolute; Exp., experimental. **d**, Heatmap of Spearman correlations between AI-predicted and experimentally measured fold changes

only from [0.16] to [0.38]. What emerges is a frustrated landscape for ligand optimization, with affinity or potency improvements counterbalanced by worsening PK. Although this trade-off has been well noted in medicinal chemistry^{40,41}, this set can quantify its systematic impact.

Prediction of analogue properties

The current benchmark only comprises several hundred analogues. If we could predict the results of the small perturbations, rather than having to experimentally test them, it could be much extended. We therefore asked how well modern methods could predict the affinity or potency and PK changes that we observed. In blinded experiments, we used free energy perturbation, arguably the highest level of theory available to the field, with the program FEP+ (Schrödinger) to predict the changes in activity between the parents and analogues (Fig. 4a–c). Similarly, we asked how well we could predict in vitro PK changes between the parents and analogues, using widely accessible tools (Fig. 4d).

There was a strong overall correlation between the FEP+-predicted and experimentally measured binding free energies (ΔG), with 188 of the 219 predictions falling within ± 2 kcal mol⁻¹ of the experimental values, and 129 within ± 1 kcal mol⁻¹, whereas the set had a mean unsigned error for ΔG of 1.06 kcal mol⁻¹ (Fig. 4a,b and Supplementary Table 9). Of the 18 parents, the FEP+ predictions on 14 had ΔG mean unsigned error below 1.2 kcal mol⁻¹ (Supplementary Table 10). If one considers the correlations qualitatively, of the 35 analogues predicted to lose affinity by 1 kcal mol⁻¹ or more by FEP+ (that is, in the range of confident prediction), 20 did so experimentally. Meanwhile, of the 34 predicted to increase in affinity or potency by more than 1 kcal mol⁻¹ by FEP+, 14 did so experimentally, whereas 12 were experimentally neutral and eight

across six in vitro PK properties, evaluated using three models. The colour intensity indicates correlation strength; correlations were assessed by two-sided Spearman rank correlation; pairwise complete-case n for Deep-PK, ADMET-AI and ADMETlab 3.0 were 218, 218 and 218, respectively, for solubility; 222 for the comparable ADMETlab 3.0 plasma-stability end point (the Deep-PK and ADMET-AI end points were not directly comparable); 222, 222 and 222, respectively, for permeability; 222, 222 and 222, respectively, for microsomal stability; 221, 221 and 221, respectively, for hERG; and 195, 195 and 195, respectively, for fraction unbound. Exact two-sided P values for all correlations are provided in the Source Data. ** $P < 0.01$ and *** $P < 0.001$. $n = 219$ analogues (**a,b**).

showed decreased potency. Categorization with a 1 kcal mol⁻¹ threshold revealed a significant association between the direction and strength of FEP+ predictions and experimentally observed affinity or potency changes ($\chi^2 = 36.4$ and $P < 0.001$; Supplementary Table 11).

Several challenges faced FEP predictions here, which future work may share. The predictive power was probably weakened by our inability to measure activity for weak analogues, reducing activity range and making it impossible to measure affinities predicted to be very low by FEP+. We used an automated workflow to generate initial poses for FEP; on inspection, most of the outliers reflected errors in these poses. FEP accuracy is impacted by the quality of the ligand-complexed structures⁴², and the predictions here began with docked poses, which are less reliable than experimental poses. For the 25 analogue structures that we determined by crystallography (for Mac1) and by cryo-electron microscopy (cryo-EM; for an α_2 analogue, below), where predicted and experimental poses closely superimposed, FEP+ energies were largely consistent with experiment (Fig. 4c).

To predict in vitro PK^{43,44}, we used ADMET-AI, ADMETlab 3.0 and Deep-PK^{45–47}. The predictions of all three correlated with the experimental measurements of plasma fraction unbound and with permeability, whereas microsomal stability, plasma stability and solubility, which are difficult to predict^{48,49}, were essentially uncorrelated. Correlations in the 0.4–0.5 range for fraction unbound and permeability may often be strong enough to guide compound selection and these open-access tools may have broad impact. Still, even with a correlation of 0.53 in fraction unbound, the highest that we observed, 28 analogues that had substantially increased fraction unbound experimentally were predicted to have lower levels than their parents, and 16 analogues that had higher protein binding than their parents experimentally were predicted to

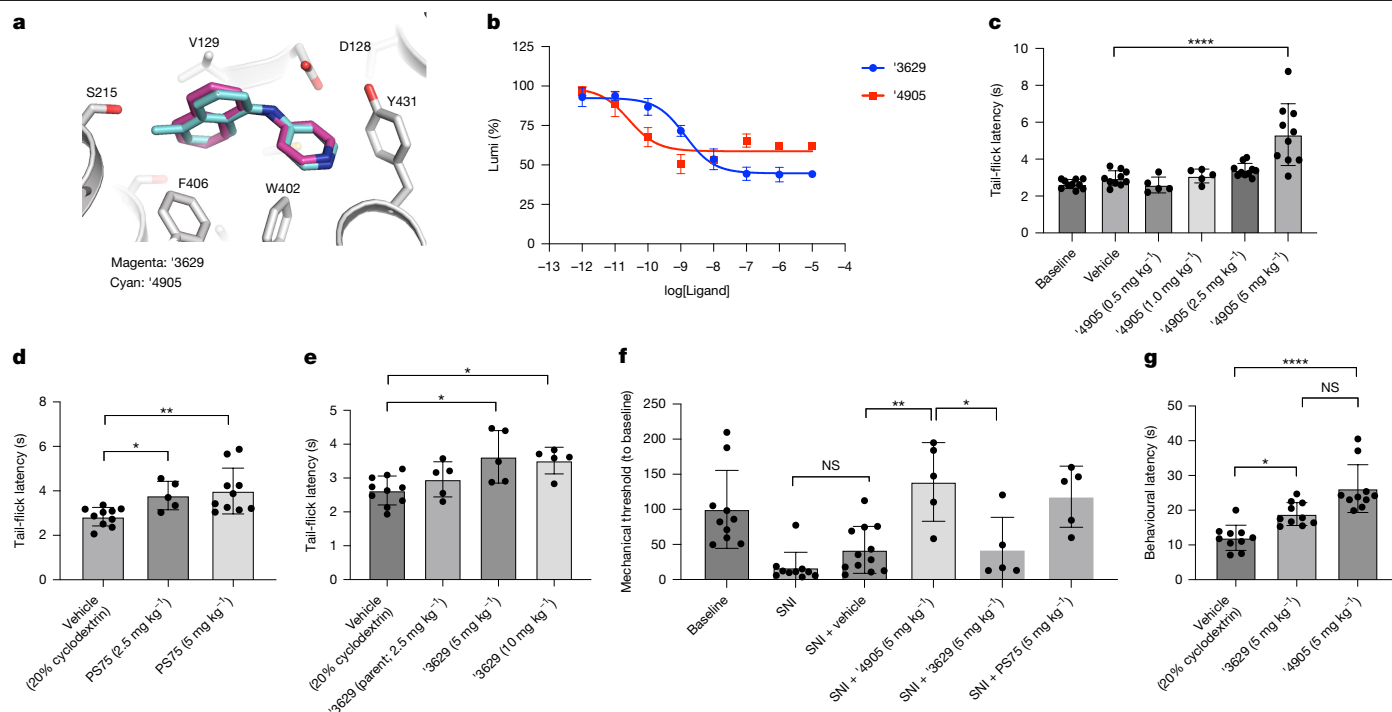


Fig. 5 | Methyl addition enhances α_2 A agonist analgesia in vivo. **a**, Docked parent '3629 superposed on the cryo-EM structure of the methyl analogue '4905. **b**, Concentration-response curves for α_2 AAR activation by '3629 and '4905; the high-concentration upturn for '4905 may reflect G_s activation. Data are mean $\pm$ s.e.m. from $n = 3$ independent experiments, each with two technical replicates. **c–e**, Tail-flick latencies after vehicle and '4905 (**c**), PS75 (**d**) or '3629 (**e**) at the indicated doses. **f**, Mechanical withdrawal threshold at baseline, after spared nerve injury (SNI), and after SNI + vehicle, '4905, '3629 or PS75 (5 mg kg⁻¹ each). **g**, Hotplate latency after vehicle, '3629 or '4905 (5 mg kg⁻¹). Mice were male C57BL/6J, 7–8 weeks of age; assignment was randomized and testing was blinded. For **c–g**, data are mean $\pm$ s.d.; points are individual mice; and group

sizes (left to right) were: $n = 10, 10, 5, 5, 10$ and 10 (**c**); $n = 10, 5$ and 10 (**d**); $n = 10, 5$ and 5 (**e**); $n = 10, 10, 12, 5, 5$ and 5 (**f**); $n = 10, 10$ and 10 (**g**). Tests were two-way analysis of variance (ANOVA) with Dunnett's test (**c**), Kruskal–Wallis with Dunn's test (**d, e**), two-way ANOVA with Tukey's test (**f**) and Friedman with Dunn's test (**g**). Exact P values: vehicle versus increasing '4905 doses of 0.9759, 0.9841, 0.7141 and <0.0001 (**c**); vehicle versus PS75 of 0.0301 and 0.0043 (**d**); vehicle versus '3629 of >0.9999 , 0.0393 and 0.0164 (**e**); SNI versus SNI + vehicle of 0.7416, SNI + vehicle versus SNI + '4905 of 0.0037, SNI + vehicle versus SNI + PS75 of 0.0333, SNI + '4905 versus SNI + '3629 of 0.0188 (**f**); vehicle versus '3629, vehicle versus '4905 and '4905 versus '3629 of 0.0417, <0.0001 and 0.2209, respectively (**g**). * $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$ and not significant (NS).

have less. Together, although the FEP and in vitro PK predictions can profitably model the effects of the small-perturbation analogues, their correlations with experiment are not yet strong enough to allow this benchmark to be extended by calculation alone.

Multi-parameter optimization for an improved analgesic

Notwithstanding the frustrated landscape for ligand affinity or potency and PK improvement, medicinal chemists can often navigate this multi-parameter landscape to improve the properties of lead molecules. Here too, there were analogues that improved in affinity or potency with only a modest decrease in PK. Among these was an analogue of an α_2 A parent, compound '4905, where a single methyl group substitution improved potency 52-fold (Fig. 5a,b). In a cryo-EM structure of the '4905– α_2 A–G protein complex determined to 2.8 Å resolution (Extended Data Table 5), '4905 superposed with the docking prediction for the parent '3629 and with the experimental structure of a previously described member of this family, '9087. This suggests that the methyl perturbation did not change the orientation of the analogue and that this family may be understood in the same structural context. A slight upwards deflection at high concentrations of '4905 in the cAMP dose–response curves may reflect secondary G_s coupling, supported by pertussis toxin-treated cAMP assays showing a residual, concentration-dependent increase in cAMP for '4905 but not '3629 (Supplementary Fig. 9a). Structurally, the added methyl group fills a hydrophobic sub-pocket (Fig. 5a and Supplementary Fig. 9b), helping to explain its improved potency.

Although the improved potency of '4905 came at a cost to its in vitro PK, the effects were modest: compared with its parent, '3629, the fraction unbound decreased by 57% and the microsomal half-life was reduced by 41% (Supplementary Tables 2–7). The in vivo PK of '4905 qualitatively reflected the in vitro results, with cerebrospinal fluid concentrations—a proxy for fraction unbound in the brain⁵⁰—reduced by 36%, whereas half-life was reduced from 196 to 15.2 min (Supplementary Table 12). Accordingly, we compared the in vivo analgesia of '4905 to that of its parent '3629 and to PS75, the most potent analgesic known for this family of α_2 A agonists⁵¹. Despite the deterioration in PK, in mouse nociception assays including tail flick, neuropathic pain (SNI) and hotplate, '4905 was far more potent than the parent '3629 and at least as potent as PS75 (Fig. 5c–g). If ligand optimization is a frustrated landscape, it remains a navigable one⁴¹. In addition to beginning to provide a random background for such efforts, this study supports a strategy by which that landscape may be reconnoitered^{23,24}.

Discussion

Although ligand optimization is at the heart of drug discovery, quantifying the success of optimization strategies has been difficult owing to a lack of a random background. Four key observations emerge from this study. First, systematic and unbiased small perturbations to 18 scaffolds revealed that 11.3% of the analogues improved tenfold or more in affinity or potency versus the parent (Table 1); 'magic methyls'^{25,52} were common in this set. These improvements were observed in five of six targets, across ligand physical properties and across five orders of magnitude of parent activity. This begins to establish a background

expectation for affinity or potency optimization against which methods and designs may be compared. It may also suggest a simple, unbiased approach for achieving substantial jumps in affinity or potency, improvements that will not always be obvious from structural analyses or simulations (Figs. 4 and 5 and Extended Data Fig. 4) but may complement them. Second, improved affinity or potency typically came at the cost of ligand PK, with plasma fraction unbound, stability to liver microsomes and permeability, among other terms, declining as analogue affinity or potency increased. The individual *in vitro* PK terms were orthogonal to each other and several were anti-correlated with affinity or potency improvement (Fig. 3). This suggests a frustrated landscape for ligand optimization. Although the qualitative challenges and trade-offs in ligand optimization are known to medicinal chemists^{8,40,41}, this study has begun to quantify them across a range of target classes and ligand properties. Third, computational prediction of affinity or potency and PK had some success anticipating the trends that we observed (Fig. 4). Still, their correlations with experiment remained loose enough to preclude replacing systematic exploration and testing of molecules as this background set is expanded, something seen with other potency prediction methods⁵³. Fourth, although potency trades off against PK, it is possible to find analogues where it rises sufficiently to overcome drops in exposure and half-life. Thus, the α_2A analogue '4905 is the most potent analgesic in a large series, and is far more potent *in vivo* than its parent, despite experiencing these trade-offs.

At first glance, this study may recall classical medicinal chemistry where conservative and ligand-based approaches such as Topliss Trees¹¹ dominated the field. However, this would be to misremember the classic age of medicinal chemistry on two counts. First, the systematic small perturbations explored here were rarely used then, partly because the building blocks and reactions to support them were lacking. The ability to systematically explore small perturbations across scaffolds reflects advances in synthetic chemistry and the new building-block approaches that have expanded readily available compounds^{54,55}. Second, the logic behind Topliss Trees and other quantitative structure–activity approaches^{11,12} was anchored in ligand physical chemistry without reference to receptor structure. For instance, in Topliss Trees the success of methyl and chloro derivatives taught opposite lessons because of their different electronic effects on the ligand, whereas the steric and non-polar properties of the two groups will seem similar from the view 'inside the receptor'^{52,56}. Meanwhile, the role of a random background would have seemed foreign. Thus, although a strategy of systematic small perturbations may smack of the 'methyl, ethyl, propyl, butyl, futile' approach with which classic medicinal chemistry has been tarred^{57,58}, it has a different basis in theory and reflects new synthetic opportunities.

Certain limitations of this study merit airing. At 257 analogues and six targets, the set remains small and inevitably, if unintentionally, biased. Expanding to additional targets and ligand classes might reveal different trends. We note that the six targets do recognize a range of chemotypes, including large polar co-factors (Mac1), large hydrophobic lipids (CB2), anionic β -lactams (AmpC), cationic neurotransmitters (α_2A and SERT) and peptides (μ OR). Experimentally, we assessed activity changes using a transporter uptake assay (SERT), radioligand competition binding assays (α_2A and CB2), second messenger assays (α_2A , CB2 and MOR), and binding and enzyme activity (Mac1 and AmpC) assays. Although ligand and competition correlates with K_i , agonist EC_{50} values are affected by receptor expression. Because we compared the relative activities of analogues and parents, controlled for receptor expression (Supplementary Fig. 4), and measured effects at low expression levels (Supplementary Fig. 5), the effects of receptor expression on relative activity should be modest, although they cannot be completely discounted. For these and related reasons, changes of EC_{50} between agonist parent and analogues cannot be read as changes in affinity the way that changes in K_i can be, although EC_{50} changes remain the relevant metric for agonist activity. Overall, the impact of

these effects may be inspected in full concentration–response curves for parent–ligand pairs with more than threefold activity improvement (Supplementary Fig. 2). In the area of *in vitro* PK, our use of liver microsomes rather than hepatocytes to measure metabolic stability meant that some types of metabolism were missed, including glucuronidation. This might make groups such as phenolic hydroxyls seem less labile than they would be *in vivo*. More broadly, for all but two molecules, we only measured *in vitro* not *in vivo* PK. Whereas the latter are widely used in ligand optimization, their prediction of *in vivo* behaviour is only approximate.

These limitations should not obscure the main observations of this study. Overall, 11.3% of systematic and unbiased small perturbations improved analogue affinity or potency tenfold or more, beginning to establish a background expectation for the likelihood of substantial ligand affinity or potency improvement and a systematic approach to doing so^{23,24}. Balancing this was a concomitant deterioration in ligand PK, which will lower the exposure and half-life of a ligand *in vivo*, counteracting the improvements in affinity or potency. Navigating this multi-parameter space is at the heart of medicinal chemistry; this study supports the development of quantitative models to do so.

Online content

Any methods, additional references, Nature Portfolio 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/s41586-026-11013-5>.

- Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *Br. J. Pharmacol.* **162**, 1239–1249 <https://doi.org/10.1111/j.1476-5381.2010.01127.x> (2011).
- Bleicher, K. H., Bohm, H. J., Muller, K. & Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* **2**, 369–378 <https://doi.org/10.1038/nrd1086> (2003).
- Gaulton, A. et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 <https://doi.org/10.1093/nar/gkr777> (2012).
- Zdrazil, B. et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* **52**, D1180–D1192 <https://doi.org/10.1093/nar/gkad1004> (2024).
- Paul, S. M. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **9**, 203–214 <https://doi.org/10.1038/nrd3078> (2010).
- Lipinski, C. A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* **1**, 337–341 <https://doi.org/10.1016/j.ddtec.2004.11.007> (2004).
- Stumpfe, D., Hu, H. & Bajorath, J. Evolving concept of activity cliffs. *ACS Omega* **4**, 14360–14368 <https://doi.org/10.1021/acsomega.9b02221> (2019).
- Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* **6**, 881–890 <https://doi.org/10.1038/nrd2445> (2007).
- Wunberg, T. et al. Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits. *Drug Discov. Today* **11**, 175–180 [https://doi.org/10.1016/S1359-6446\(05\)03700-1](https://doi.org/10.1016/S1359-6446(05)03700-1) (2006).
- Jorgensen, W. L. Efficient drug lead discovery and optimization. *Acc. Chem. Res.* **42**, 724–733 <https://doi.org/10.1021/ar800236t> (2009).
- Topliss, J. G. Utilization of operational schemes for analog synthesis in drug design. *J. Med. Chem.* **15**, 1006–1011 <https://doi.org/10.1021/jm00280a002> (1972).
- Hansch, C. & Fujita, T. ρ - σ - π Analysis: a method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* **86**, 1616–1626 (1964).
- Dong, J. et al. ADMETLab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *J. Cheminform.* **10**, 29 <https://doi.org/10.1186/s13321-018-0283-x> (2018).
- Wang, L. et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **137**, 2695–2703 <https://doi.org/10.1021/ja512751q> (2015).
- Klimovich, P. V., Shirts, M. R. & Mobley, D. L. Guidelines for the analysis of free energy calculations. *J. Comput. Aided Mol. Des.* **29**, 397–411 <https://doi.org/10.1007/s10822-015-9840-9> (2015).
- Hajduk, P. J. & Sauer, D. R. Statistical analysis of the effects of common chemical substituents on ligand potency. *J. Med. Chem.* **51**, 553–564 <https://doi.org/10.1021/jm070838y> (2008).
- Kramer, C., Kallioikoski, T., Gedeck, P. & Vulpetti, A. The experimental uncertainty of heterogeneous public data. *J. Med. Chem.* **55**, 5165–5173 <https://doi.org/10.1021/jm300131x> (2012).
- Booker, T. R., Jackson, B. C. & Keightley, P. D. Detecting positive selection in the genome. *BMC Biol.* **15**, 98 <https://doi.org/10.1186/s12915-017-0434-y> (2017).

19. Pritchard, J. K. & Cox, N. J. The allelic architecture of human disease genes: common disease—common variant... or not? *Hum. Mol. Genet.* **11**, 2417–2423 <https://doi.org/10.1093/hmg/11.20.2417> (2002).
20. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 <https://doi.org/10.1038/nature08494> (2009).
21. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11**, 801–807 <https://doi.org/10.1038/Nmeth.3027> (2014).
22. Wells, J. A. Systematic mutational analyses of protein protein interfaces. *Method Enzymol.* **202**, 390–411 (1991).
23. Pennington, L. D., Aquila, B. M., Choi, Y., Valiulin, R. A. & Muegge, I. Positional analogue scanning: an effective strategy for multiparameter optimization in drug design. *J. Med. Chem.* **63**, 8956–8976 <https://doi.org/10.1021/acs.jmedchem.9b02092> (2020).
24. Muegge, I. & Hu, Y. In silico positional analogue scanning with Amber GPU-TI. *J. Chem. Inf. Model.* **62**, 4448–4459 <https://doi.org/10.1021/acs.jcim.2c00860> (2022).
25. Schönher, H. & Cernak, T. Profound methyl effects in drug discovery and a call for new C–H methylation reactions. *Angew. Chem. Int. Ed.* **52**, 12256–12267 (2013).
26. Cramer, J., Sager, C. P. & Ernst, B. Hydroxyl groups in synthetic and natural-product-derived therapeutics: a perspective on a common functional group. *J. Med. Chem.* **62**, 8915–8930 (2019).
27. Chiodi, D. & Ishihara, Y. “Magic chloro”: profound effects of the chlorine atom in drug discovery. *J. Med. Chem.* **66**, 5305–5331 (2023).
28. Gillis, E. P., Eastman, K. J., Hill, M. D., Donnelly, D. J. & Meanwell, N. A. Applications of fluorine in medicinal chemistry. *J. Med. Chem.* **58**, 8315–8359 (2015).
29. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* **9**, 430–431 [https://doi.org/10.1016/S1359-6446\(04\)03069-7](https://doi.org/10.1016/S1359-6446(04)03069-7) (2004).
30. Zhu, T. et al. Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* **56**, 6560–6572 <https://doi.org/10.1021/jm301916b> (2013).
31. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. The maximal affinity of ligands. *Proc. Natl Acad. Sci. USA* **96**, 9997–10002 <https://doi.org/10.1073/pnas.96.18.9997> (1999).
32. Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **49**, 377–389 <https://doi.org/10.1021/ci800324m> (2009).
33. Thomas, M., Bender, A. & de Graaf, C. Integrating structure-based approaches in generative molecular design. *Curr. Opin. Struct. Biol.* **79**, 102559 <https://doi.org/10.1016/j.sbi.2023.102559> (2023).
34. Flowers, J. et al. Expanding automated multiconformer ligand modeling to macrocycles and fragments. *eLife* **14**, RP103797 <https://doi.org/10.7554/eLife.103797> (2025).
35. Di, L., Kerns, E. H. & Carter, G. T. Drug-like property concepts in pharmaceutical design. *Curr. Pharm. Des.* **15**, 2184–2194 <https://doi.org/10.2174/138161209788682479> (2009).
36. Morgan, P. et al. Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving phase II survival. *Drug Discov. Today* **17**, 419–424 <https://doi.org/10.1016/j.drudis.2011.12.020> (2012).
37. Houston, J. B. Utility of in vitro drug metabolism data in predicting in vivo metabolic clearance. *Biochem. Pharmacol.* **47**, 1469–1479 [https://doi.org/10.1016/0006-2952\(94\)90520-7](https://doi.org/10.1016/0006-2952(94)90520-7) (1994).
38. Iwatsubo, T. et al. Prediction of in vivo drug metabolism in the human liver from in vitro metabolism data. *Pharmacol. Ther.* **73**, 147–171 [https://doi.org/10.1016/s0163-7258\(96\)00184-2](https://doi.org/10.1016/s0163-7258(96)00184-2) (1997).
39. Amidon, G. L., Lennernas, H., Shah, V. P. & Crison, J. R. A theoretical basis for a biopharmaceutical drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm. Res.* **12**, 413–420 <https://doi.org/10.1023/a:1016212804288> (1995).
40. Hopkins, A. L., Keseru, G. M., Leeson, P. D., Rees, D. C. & Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **13**, 105–121 <https://doi.org/10.1038/nrd4163> (2014).
41. Young, R. J. & Leeson, P. D. Mapping the efficiency and physicochemical trajectories of successful optimizations. *J. Med. Chem.* **61**, 6421–6467 <https://doi.org/10.1021/acs.jmedchem.8b00180> (2018).
42. Rocklin, G. J., Mobley, D. L., Dill, K. A. & Hunenberger, P. H. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. *J. Chem. Phys.* **139**, 184103 <https://doi.org/10.1063/1.4826261> (2013).
43. Kim, H. et al. Artificial intelligence in drug discovery: a comprehensive review of data-driven and machine learning approaches. *Biotechnol. Bioprocess Eng.* **25**, 895–930 <https://doi.org/10.1007/s12257-020-0049-y> (2020).
44. Kumar, A., Kini, S. G. & Rathi, E. A recent appraisal of artificial intelligence and in silico ADMET prediction in the early stages of drug discovery. *Mini Rev. Med. Chem.* **21**, 2788–2800 <https://doi.org/10.2174/1389557521666210401091147> (2021).
45. Fu, L. et al. ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support. *Nucleic Acids Res.* **52**, W422–W431 <https://doi.org/10.1093/nar/gkac236> (2024).
46. Myung, Y., de Sa, A. G. C. & Ascher, D. B. Deep-PK: deep learning for small molecule pharmacokinetic and toxicity prediction. *Nucleic Acids Res.* **52**, W469–W475 <https://doi.org/10.1093/nar/gkac254> (2024).
47. Swanson, K. et al. ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btae416> (2024).
48. Bergstrom, C. A., Wassvik, C. M., Johansson, K. & Hubatsch, I. Poorly soluble marketed drugs display solvation limited solubility. *J. Med. Chem.* **50**, 5858–5862 <https://doi.org/10.1021/jm0706416> (2007).
49. Obach, R. S. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. *Drug Metab. Dispos.* **27**, 1350–1359 (1999).
50. Friden, M. et al. Structure-brain exposure relationships in rat and human using a novel data set of unbound drug concentrations in brain interstitial and cerebrospinal fluids. *J. Med. Chem.* **52**, 6233–6243 <https://doi.org/10.1021/jm901036q> (2009).
51. Fink, E. A. et al. Structure-based discovery of nonopioid analgesics acting through the α_{2A} -adrenergic receptor. *Science* **377**, eabn7065 <https://doi.org/10.1126/science.abn7065> (2022).
52. Leung, C. S., Leung, S. S., Tirado-Rives, J. & Jorgensen, W. L. Methyl effects on protein-ligand binding. *J. Med. Chem.* **55**, 4489–4500 <https://doi.org/10.1021/jm3003697> (2012).
53. Sindt, F., Bret, G. & Rognan, D. On the difficulty to rescore hits from ultralarge docking screens. *J. Chem. Inf. Model.* **65**, 5553–5566 <https://doi.org/10.1021/acs.jcim.5c00730> (2025).
54. Lyu, J. et al. Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 <https://doi.org/10.1038/s41586-019-0917-9> (2019).
55. Moroz, Y., Chuprina, A. & Mykytenko, D. Enamine REAL DataBase – an instrumental and practical vehicle for charting new regions of the relevant drug discovery chemical space. In *251st American Chemical Society National Meeting & Exposition* (ACS, 2016).
56. Verteramo, M. L. et al. Interplay of halogen bonding and solvation in protein-ligand binding. *iScience* **27**, 109636 <https://doi.org/10.1016/j.isci.2024.109636> (2024).
57. Krimmer, S. G., Betz, M., Heine, A. & Klebe, G. Methyl, ethyl, propyl, butyl: futile but not for water, as the correlation of structure and thermodynamic signature shows in a congeneric series of thermolysin inhibitors. *ChemMedChem* **9**, 833–846 <https://doi.org/10.1002/cmde.201400013> (2014).
58. Wermuth, C. G. *The Practice of Medicinal Chemistry* (Elsevier Science & Technology, 2008).

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



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

© The Author(s) 2026

Methods

ChEMBL database molecular pairs

From the PostgreSQL version of ChEMBL34 (ref. 4) (<https://doi.org/10.6019/CHEMBL.database.34>), we filtered the database for compounds with activities reported against a single protein, with either K_i , K_d , IC_{50} or EC_{50} as the activity type. To be considered a molecular pair, compounds had to have the same assay ID, the same target ID, the same activity type and the same reference document (publication) ID, in addition to having a parent–analogue relationship as defined in this work, meaning a C–H group replaced by C–OH, C–F, C–Cl, C–Br, C–CH₃ or an aromatic carbon replaced by a nitrogen. This left us with 191,732 parent–analogue pairs.

Analogue enumeration and synthesis

Eighteen parent compounds were selected from previously published literature^{51,59,60} or datasets (<https://asapdiscovery.org/outputs/molecules/#ASAP-SARS-COV-2-NSP3-MAC1>) based on their known binding activity, structural relevance or representation of diverse chemical scaffolds. Starting from a parent compound, we used RDKit (<http://www.rdkit.org/>) to identify all C–H bonds and iteratively replaced the hydrogen atom with a methyl, hydroxyl, fluoro, chloro or bromo group. We also identified aromatic carbons with two heavy-atom neighbours and replaced them with nitrogen. Every analogue generated was represented as a canonical isomeric SMILES string and added to a set to remove duplicates. For each of the 18 parents, the full set of possible analogues that could be synthesized for less than US\$400 for 10 mg was ordered.

Confidence intervals and statistics

When reported, 95% confidence intervals were derived from bootstrap resampling with 10,000 iterations. In cases in which the observed frequency of success is exactly 0, bootstrap will fail to give an upper bound for the interval. In such cases, we used the Clopper–Pearson method to estimate an upper bound based on the sample size⁶¹. Pearson and Spearman correlations with their associated P values were computed using the `scipy.stats` module from the SciPy package⁶².

FEP simulations

These were conducted using FEP+ within the Schrödinger software suite (v2025-2) with the OPLS4 force field⁶³ and the modified simple point charge water model. The default setting was used for the number of lambda windows selection where it depends on the type of perturbations; charge-changing, core hopping and all other perturbations have 24, 16, and 12 lambda windows, respectively. For alchemical transformations with charge changes, the total charge of the simulation box was kept constant by transmuting a Na⁺ or Cl[−] ion to water or vice versa. In addition, a 0.15 M concentration of NaCl was added to the simulation box of charge-changing perturbations. For α_2A , CB2 and SERT, the FEP+ membrane protocol was applied where a POPC membrane was added to the system in simulation. All other settings were kept default except that the simulation time was extended from 5 ns to 10 ns.

The default FEP map generation protocol was used with the parent compound selected as the biased node. In preparing proteins and ligands for FEP+, the Schrödinger protein preparation workflow and LigPrep were used. The initial binding poses of parent compounds were from poses generated by DOCK3.8, and analogues were aligned to the parents with severe steric clashes resolved using the FEP+ Pose Builder workflow.

Docking

Whereas no structural information was used in the design of analogues from parent compounds, for FEP+ calculations and for post hoc structural analysis, we generated ligand-bound complexes of the

parents and relevant ligands using DOCK3.8 (refs. 54,64). Ligands were docked into the receptor-binding site using grids prepared in previous studies^{51,59,60,65,66}.

Assay selection

We used one consistent readout per parent series (no mixing within a series). SERT and AmpC are reported as K_i ; Mac1 is reported as IC_{50} from peptide displacement (a scalable functional hydrolysis assay is unavailable); GPCR agonist series are reported as EC_{50} and antagonist series as K_i , such that μOR is EC_{50} only, whereas α_2A and CB2 include EC_{50} and K_i depending on the parent series. Representative concentration–response curves and the corresponding Z values for each assay are provided in Supplementary Fig. 3.

Transporter assays for SERT K_i

SERT activity was measured using the Neurotransmitter Transporter Uptake Assay Kit from Molecular Devices (#R8174), following the manufacturer's protocol with slight modifications as previously described⁶⁵. HEK293T cells (ATCC CRL-3216) stably expressing human SERT were plated in poly-L-lysine-coated 384-well black, clear-bottom plates at a density of 15,000 cells in 40 μ l per well, using DMEM supplemented with 1% dialysed FBS. Cells were incubated overnight at 37 °C with 5% CO₂ to allow adherence and recovery. The following day, the medium was removed by flicking, and cells were incubated with 25 μ l per well of test compound solutions prepared in assay buffer (1× HBSS, 20 mM HEPES, pH 7.4, supplemented with 1 mg ml^{−1} BSA) for 30 min at 37 °C. After drug treatment, 25 μ l per well of dye solution (as provided in the kit) was added directly to the wells, followed by an additional 30-min incubation at 37 °C. Fluoxetine (10 μ M) was used as a positive control for SERT inhibition. Fluorescence was measured using the FlexStation II microplate reader with excitation at 440 nm and emission at 520 nm. Relative fluorescence units were exported and analysed using GraphPad Prism 10.0 to calculate IC_{50} values, from which K_i values were derived using the Cheng–Prusoff equation.

CB2 radioligand-binding assay

CB2 receptor-binding assays were performed using membrane preparations from HEK293 cells (ATCC CRL-1573) stably expressing human CB2, following previously published methods^{67,68}. Membranes were resuspended in TME buffer containing 0.1% BSA (w/v) and 25 μ g of membrane protein was added per well. The assay was conducted using the radioligand [³H]CP-55,940 at a final concentration of 0.75 nM, prepared in assay buffer. Nonspecific binding was defined in the presence of 5 μ M unlabelled CP-55,940. Test compounds were applied at increasing concentrations to assess competition. Reactions were incubated at 30 °C for 1 h with gentle shaking. After incubation, samples were transferred to Unifilter GF/B 96-well filter plates and filtered using a Packard Filtermate-196 cell harvester (PerkinElmer). Plates were washed four times with ice-cold wash buffer (50 mM Tris-HCl, 5 mM MgCl₂ and 0.5% BSA, pH 7.4). Radioactivity bound to the filters was quantified via liquid scintillation counting. Specific binding was calculated by subtracting nonspecific binding from total binding. IC_{50} and K_i values were calculated using nonlinear regression in GraphPad Prism 9 using the Cheng–Prusoff equation.

α_2A receptor-binding assay

α_2A binding was performed using membrane preparations from insect cells (Expression Systems, 94-001S) expressing human α_2A receptors, as previously described⁵¹. Membranes were incubated with increasing concentrations of test compounds and 5 nM [³H]rauwolscine in buffer containing 20 mM HEPES (pH 7.5) and 100 mM NaCl, at room temperature for 2 h. After incubation, samples were filtered onto GF/B filter plates, washed with ice-cold buffer and radioactivity was quantified by liquid scintillation counting. IC_{50} and K_i values were derived using nonlinear regression in GraphPad Prism.

GloSensor cAMP assay for α_2 A, CB2 and MOR

The GloSensor cAMP assay was performed following the manufacturer's instructions (Promega) with slight modifications as previously described^{51,69}. In brief, wild-type human α_2 A, CB2 and μ OR were cloned into the pcDNA3.1 vector and co-transfected with the 22F cAMP GloSensor plasmid into HEK293T cells (ATCC CRL-3216) cultured in six-well plates. After 24 h, cells were reseeded into 96-well white plates in CO₂-independent medium and equilibrated with GloSensor cAMP reagent as per the manufacturer's protocol. Cells were incubated for 1 h at 37 °C followed by 1 h at room temperature. Where applicable, 10 μ M forskolin was used to elevate basal cAMP levels for assessing receptor-mediated inhibition. Serially diluted test compounds were added, and luminescence signals were recorded using a PerkinElmer microplate reader. Data were analysed using GraphPad Prism 9.0 to calculate EC₅₀ or IC₅₀ values.

AmpC β -lactamase inhibition assay

The AmpC β -lactamase inhibition assay was performed as previously described⁶⁰. The candidate inhibitors were dissolved in DMSO (20 mM stock) and diluted to maintain a constant 1% DMSO (v/v) in 50 mM sodium cacodylate buffer (pH 6.5). Assays were performed in the presence of 0.01% Triton X-100 to reduce aggregation artefacts. AmpC enzymatic activity was monitored spectrophotometrically using CENTA or nitrocefin as substrates. Initial screening was performed at 200 μ M, 100 μ M and 40 μ M compound concentrations. Substrate concentrations were selected based on known K_m values to achieve defined [S]: K_m ratios: for CENTA ([S] = 50 μ M, K_m = 27.6 μ M) and nitrocefin ([S] = 100 μ M or 28 μ M, K_m = 180 μ M). Reactions were carried out in 96-well format on a BMG Labtech CLARIOstar plate reader, with substrate and enzyme injected into wells containing the inhibitor, followed by kinetic measurement over 50 s. IC₅₀ values were determined by fitting inhibition curves in GraphPad Prism using a fixed Hill coefficient of 1, and K_i values were calculated using the Cheng–Prusoff equation⁷⁰.

HTRF assay for Mac1

Binding of the compounds to Mac1 was assessed by the displacement of an ADPr-conjugated biotin peptide from His6-tagged protein using a homogeneous time-resolved fluorescence (HTRF)-based assay, as previously described⁷¹. The expression sequences used for SARS-CoV-2 Mac1 are listed below. All proteins were expressed and purified as previously described for SARS-CoV-2 Mac1 (ref. 71). Compounds were dispensed into ProxiPlate-384 Plus (PerkinElmer) assay plates using an Echo 650 Liquid Handler (Beckman Coulter). Binding assays were conducted in a final volume of 16 μ l with 12.5 nM NSP3 Mac1 protein, 200 nM peptide ARTK(Bio)QTARK(Aoa- RADP)S (Cambridge Peptides), 1:20,000 anti-His6-Eu3⁺ cryptate (HTRF donor; PerkinElmer AD0402) and 1:500 streptavidin-XL665 (HTRF acceptor; PerkinElmer 610SAXLB) in assay buffer (25 mM HEPES pH 7.0, 20 mM NaCl, 0.05% bovine serum albumin and 0.05% Tween-20, the latter also to reduce aggregation artefacts). Assay reagents were dispensed manually into plates using an electronic multichannel pipette. Mac1 and peptide were pre-incubated for 30 min at room temperature before HTRF reagents were added. Fluorescence was measured after a 1-h incubation at room temperature using a Perkin Elmer EnVision 2105-0010 Dual Detector Multimode microplate reader with dual emission protocol (A = excitation of 320 nm and emission of 665 nm, and B = excitation of 320 nm and emission of 620 nm). Compounds were tested in triplicate in a 14-point dose response. Raw data were processed to give an HTRF ratio (channel A/B $\times$ 10,000), which was used to generate IC₅₀ curves using nonlinear regression using GraphPad Prism v10.0.2 (GraphPad Software).

hERG channel inhibition

hERG channel inhibition was evaluated using a Thallium Flux assay on HEK293 cells (ATCC CRL-1573) stably expressing the hERG potassium

channel. Cells were seeded at a density of 8,000 cells per well in 384-well poly-D-lysine-coated plates and incubated for 24 h under standard conditions (37 °C at 5% CO₂). The next day, a thallium-sensitive dye was added to the cells, followed by a 1-h incubation to ensure dye uptake. Test compounds were added to achieve a final concentration of 30 μ M in 0.5% DMSO, and the cells were incubated for an additional 30 min at room temperature. Subsequently, a stimulation buffer containing thallium was added, and fluorescence measurements were taken using a FLIPR Tetra system. Data were collected every 3 s for 3 min (excitation of 470–495 nm and emission of 515–575 nm). Fluorescence intensity over time was analysed to calculate the area under the curve from which percentage inhibition was determined versus haloperidol at 100 μ M (positive control) and vehicle (DMSO).

α_2 A receptor purification and structure determination

Wild-type human α_2 AAR was cloned into a pVL1392 vector with an N-terminal FLAG tag. The construct was expressed in *Spodoptera frugiperda* (Sf9) insect cells using the BestBac system. Cells at a density of 4×10^6 cells per millilitre were infected with virus and incubated for 48 h at 27 °C. The receptor was solubilized and purified by FLAG affinity chromatography and size-exclusion chromatography in the presence of 10 μ M compound '4905. Monomeric peak fractions were concentrated and used for G protein complex formation. G_{ao}G _{β 1 γ 2} heterotrimeric G proteins were expressed in Hi5 insect cells (Expression Systems, 94-002S) and purified using Ni²⁺ affinity following detergent solubilization and dephosphorylation. The final α_2 AAR–G_{ao}G _{β 1 γ 2}–scFv16 complex⁷² was assembled in the presence of '4905 and purified by size-exclusion chromatography. Cryo-EM grids were prepared using UltrAufoil R1.2/1.3 300-mesh grids and vitrified in liquid ethane. Data were collected on a Titan Krios G3 electron microscope equipped with a K3 direct electron detector. Image processing was performed using cryoSPARC, yielding a final reconstruction at approximately 2.8 Å resolution. Model building and refinement were performed with Protein Data Bank (PDB) 7EJ8 as a starting model using ChimeraX⁷³, Phenix⁷⁴ and Coot⁷⁵. The final structures have been deposited in the PDB with accession codes 9PLO and 9PLN. Structural representations of protein–ligand complexes were generated using the PyMOL Molecular Graphics System v2.5.5 (Schrödinger).

Mac1 purification and crystallization

Wild-type Mac1 protein (P4₃ construct, residues 3–169) was expressed in *Escherichia coli* BL21(DE3) as an N-terminal His₆-tagged construct and purified by Ni²⁺-affinity chromatography⁷¹. The His-tag was cleaved with TEV protease, followed by size-exclusion chromatography (Superdex 75) in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM dithiothreitol. Purified protein was concentrated to 40 mg ml^{−1} for crystallization and stored at −80 °C.

Crystals were obtained by sitting-drop vapour diffusion in 28% PEG 3000 and 100 mM CHES (pH 9.5). Compounds (100 mM in DMSO) were added to crystal drops using an Echo 650 acoustic dispenser⁷⁶ to a final concentration of 10 mM. Crystals were incubated at room temperature for 2–4 h and vitrified in liquid nitrogen without additional cryoprotection. X-ray diffraction data were collected at 100 K at the Advanced Light Source (beamline 8.3.1) using an X-ray wavelength of 0.88557 Å, and diffraction data were processed using XDS⁷⁷ and Aimless⁷⁸. Structures were determined to resolutions ranging from 0.97 to 1.02 Å. Ligands with low occupancy or conformational disorder were modelled using PanDDA⁷⁹ and Coot⁷⁵ and refined using phenix.refine⁸⁰ as previously described⁸¹. The structure of '3184 was determined using a racemic preparation of the compound ('9037). X-ray data collection and refinement statistics are shown in Extended Data Tables 1–4, as are the 32 PDB IDs. Structural representations of protein–ligand complexes were generated using the PyMOL Molecular Graphics System v2.5.5 (Schrödinger).

Microsomal stability

Microsomal stability of compounds was evaluated using pooled mouse liver microsomes (M3000/lot #2010026, XenoTech) to estimate their metabolic stability and predict hepatic clearance. Each compound was incubated at 2 μM in a reaction mixture containing 0.42 mg ml⁻¹ microsomal protein, phosphate buffer (100 mM, pH 7.4), MgCl₂ (3.3 mM), NADPH (3 mM), glucose-6-phosphate (5.3 mM) and glucose-6-phosphate dehydrogenase (0.67 units per millilitre). Reactions were conducted at 37 °C in 96-well plates with shaking at 100 rpm. Samples were collected at five time points (0, 7, 15, 25 and 40 min), and reactions were quenched by adding five volumes of acetonitrile containing an internal standard. After centrifugation at 5,500 rpm for 5 min, the supernatants were analysed via high-performance liquid chromatography–tandem mass spectrometry (HPLC–MS/MS). The elimination rate constant (kel), half-life ($t_{1/2}$) and intrinsic clearance (Clint) were calculated by plotting the natural logarithm of the remaining parent compound versus time. Stability was compared with reference standards such as imipramine and propranolol.

Plasma protein binding

Plasma protein binding was measured using equilibrium dialysis with a 14-kDa molecular weight cut-off membrane in a 96-well HTD96b dialyser. Mouse plasma containing 1 μM test compound (0.005% DMSO and 1% acetonitrile) was placed in one chamber, and phosphate-buffered saline (PBS, pH 7.4) in the opposing chamber. The assembled plates were incubated at 37 °C with 5% CO₂ and approximately 95% humidity, shaking at 250 rpm for 5 h to reach equilibrium. After incubation, aliquots from each chamber were mixed with equal volumes of the blank opposite matrix and processed with acetonitrile containing internal standard. Supernatants obtained after centrifugation were analysed by HPLC–MS/MS. The percentage of compound bound to plasma proteins was calculated using the peak area ratio in buffer to plasma compartments. Recovery and stability standards were included to ensure accuracy and reliability. Verapamil served as a reference control. Most compounds showed moderate binding, typically ranging between 75% and 85%.

Plasma stability

Plasma stability was assessed in non-sterile mouse plasma (Li-heparin treated) at 1 μM concentration (final DMSO content was 0.005%). Incubations were carried out in aliquots of 60 μl each (two per time point) at 37 °C under 5% CO₂ and high humidity (approximately 95%). The reactions were quenched with 240 μl of 90% acetonitrile containing an internal standard at 0, 20, 40, 60 and 120 min, followed by centrifugation at 6,000 rpm for 5 min. Supernatants were analysed via HPLC–MS/MS to determine the percentage of parent compound remaining at each time point. Data were plotted to calculate compound half-lives ($t_{1/2}$). Reference compounds, verapamil and propantheline, were used as high and low stability controls, respectively. This assay is critical for identifying compounds susceptible to degradation by plasma esterases or hydrolytic enzymes, helping inform PK optimization strategies during lead selection.

Thermodynamic solubility

Aqueous thermodynamic solubility was determined in PBS (pH 7.4) for 247 compounds using a shake-flask method followed by UV absorbance quantification. Dry powder compounds were dissolved in PBS to a theoretical concentration of 4 mM and incubated in duplicate at 25 °C for 4 h and 24 h with shaking. After incubation, samples were filtered using HTS 96-well filter plates. The filtrates were diluted twofold in acetonitrile with 4% DMSO for UV analysis. The incubation samples for charged molecules were additionally diluted tenfold with 50% acetonitrile–PBS with 2% final DMSO. Calibration curves (0–200 μM) were prepared in 50% acetonitrile–PBS (2% final DMSO). Absorbance was

measured between 230 and 550 nm using a SpectraMax Plus microplate reader. Compound-specific absorbance maxima were used to calculate concentration using SoftMax Pro and Excel. The assay dynamic range was approximately 2–400 μM (about 20–4,000 μM for charged molecules), with values near the upper limit treated as semi-quantitative. Ondansetron was used as a reference compound. This method reflects equilibrium solubility under physiologically relevant conditions and helps to rank compounds for formulation feasibility.

PAMPA-BBB

Passive blood–brain barrier (BBB) permeability was estimated using PAMPA (PAMPA-BBB) with a phospholipid-coated membrane simulating the brain endothelium. Test compounds (50 μM in Prisma HT buffer, pH 7.4, with 0.5% DMSO) were added to donor wells, whereas brain sink buffer was added to the acceptor wells. The donor and acceptor chambers were separated by a 0.45- μm filter membrane coated with brain polar lipids. Plates were incubated without agitation at room temperature for 4 h. Post-incubation, samples from both chambers, as well as a standard solution, were diluted with acetonitrile containing an internal standard. Apparent permeability coefficients (log Papp) were calculated based on peak area ratio. Clozapine and chlorpromazine were used as high-permeability controls, whereas ranitidine represented low permeability. The assay provides a high-throughput, non-cell-based method for estimating central nervous system exposure potential.

Behavioural analyses

Male C57BL/6J mice 7–8 weeks of age were obtained from The Jackson Laboratory (JAX strain #000664) and housed five per cage under a standard 12–12-h light–dark cycle at 20.7 °C (69.3 °F) and 58% humidity. All animal procedures were approved by the University of California, San Francisco Institutional Animal Care and Use Committee (protocol #AN208219). Animals were randomly assigned to treatment and control groups. For behavioural experiments, mice were initially placed together in a cage and allowed to move freely for a few minutes; each mouse was then randomly selected, injected with compound or vehicle, and placed in a separate cylinder before testing. The experimenter assessing behaviour was different from the experimenter administering injections and placing mice into the cylinders; experimenters were therefore blinded to treatment. Behavioural experiments were conducted in two independent cohorts. Mice were habituated individually in Plexiglas enclosures for 1 h before testing. Compounds were administered subcutaneously 30 min before behavioural assessment, and, where applicable, the α_2 AAR antagonist atipamezole (2 mg kg⁻¹, intraperitoneally) was given 15 min before compound injection. Tail-flick latency was measured by immersing the distal third of the tail in a 50 °C water bath and recording the withdrawal time. For the neuropathic pain model, SNI was performed under isoflurane anaesthesia by ligating and transecting two of the three branches of the sciatic nerve, sparing the sural nerve. Mechanical thresholds were assessed 7–14 days post-surgery using von Frey filaments and the up–down method, and values were normalized to the baseline of each animal. Thermal nociception was evaluated using a 55 °C hotplate, and the latency to nocifensive behaviour (paw lick or jump) was recorded with a cut-off of 45 s to prevent tissue damage. Behavioural data are summarized as group size (n), mean and s.d. (Supplementary Table 13). Statistical analyses were performed in GraphPad Prism. For the SNI experiment (Fig. 5f), we used two-way ANOVA followed by Tukey's multiple-comparisons test. The hotplate assay (Fig. 5g) was analysed using the Friedman test followed by Dunn's multiple-comparisons test. For the tail-flick assays, the 4905 dose–response (Fig. 5c) was analysed by two-way ANOVA followed by Dunnett's multiple-comparisons test against vehicle (control), whereas the PS75 and 3629 dose groups (Fig. 5d,e) were analysed using the Kruskal–Wallis test followed by Dunn's multiple-comparisons test against vehicle. The Kruskal–Wallis and Friedman tests were nonparametric; Prism calculated two-sided, multiplicity-adjusted P values for

the Dunn comparisons (Supplementary Table 13). For sample size, we did not perform formal a priori power calculations. Instead, group sizes were guided by our previous experience with these assays and by precedent in the literature. This approach may limit sensitivity to small effects. Raw data for the animal assays underlying Fig. 5c–g are provided in the Source Data file.

Reporting summary

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

Data availability

All primary data are available through this article and its Extended Data and Supplementary Information. The cryo-EM structures and maps have been deposited in the PDB and Electron Microscopy Data Bank, respectively, under the accession codes 9PLO and EMD-71719 (the '4905-bound α_2 AAR complex with G proteins and scFv16), and PDB 9PLN and EMD-71718 (the '4905-bound locally refined α_2 AAR structure). The Mac1 crystal structures and supporting structure factors have been deposited in the PDB under the accession codes 14AB, 7IIW, 7IIX, 7IIY, 7IIZ, 7IJ0, 7IJ1, 7IJ2, 7IJ3, 7IJ4, 7IJ5, 7IJ6, 7IJ7, 7IJ8, 7IJ9, 7IJA, 7IJB, 7IJC, 7IJD, 7IJE, 7IJF, 7IJG, 7IJH, 7IJI, 7IJJ, 7IJK, 7IJL, 14AM, 14AN, 14AO, 14AP and 7IJM. Source data are provided with this paper.

Code availability

ZINC tools used in the selection of the small-perturbation analogues are openly available (<https://zinc22.docking.org>).

59. Gahbauer, S. et al. Iterative computational design and crystallographic screening identifies potent inhibitors targeting the Nsp3 macrodomain of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **120**, e2212931120 (2023).
60. Liu, F. et al. The impact of library size and scale of testing on virtual screening. *Nat. Chem. Biol.* **21**, 1039–1045 <https://doi.org/10.1038/s41589-024-01797-w> (2025).
61. Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413 (1934).
62. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
63. Lu, C. et al. OPLS4: improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* **17**, 4291–4300 <https://doi.org/10.1021/acs.jctc.1c00302> (2021).
64. Balias, T. E. et al. Testing inhomogeneous solvation theory in structure-based ligand discovery. *Proc. Natl Acad. Sci. USA* **114**, E6839–E6846 (2017).
65. Singh, I. et al. Structure-based discovery of conformationally selective inhibitors of the serotonin transporter. *Cell* **186**, 2160–2175.e17 (2023).
66. Vigneron, S. F. et al. Docking 14 million virtual isoquinolines against the μ and κ opioid receptors reveals dual antagonists—inverse agonists with reduced withdrawal effects. *ACS Central Sci.* **11**, 770–790 (2025).
67. Hua, T. et al. Activation and signaling mechanism revealed by cannabinoid receptor-Gi complex structures. *Cell* **180**, 655–665.e18 (2020).
68. Hua, T. et al. Crystal structures of agonist-bound human cannabinoid receptor CB1. *Nature* **547**, 468–471 (2017).
69. Manglik, A. et al. Structure-based discovery of opioid analgesics with reduced side effects. *Nature* **537**, 185–190 (2016).
70. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
71. 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 <https://doi.org/10.1126/sciadv.abf8711> (2021).
72. Maeda, S. et al. Development of an antibody fragment that stabilizes GPCR/G-protein complexes. *Nat. Commun.* **9**, 3712 <https://doi.org/10.1038/s41467-018-06002-w> (2018).
73. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 <https://doi.org/10.1002/pro.3235> (2018).
74. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 <https://doi.org/10.1107/S2059798319011471> (2019).
75. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 <https://doi.org/10.1107/S09074444190007493> (2010).
76. Collins, P. M. et al. Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr. D Struct. Biol.* **73**, 246–255 <https://doi.org/10.1107/S205979831700331X> (2017).
77. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 <https://doi.org/10.1107/S0907444909047337> (2010).
78. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 <https://doi.org/10.1107/S0907444913000061> (2013).
79. Pearce, N. M. et al. A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 <https://doi.org/10.1038/ncomms15123> (2017).
80. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 <https://doi.org/10.1107/S0907444912001308> (2012).
81. Correy, G. J. et al. Exploration of structure-activity relationships for the SARS-CoV-2 macrodomain from shape-based fragment linking and active learning. *Sci. Adv.* **11**, eads7187 <https://doi.org/10.1126/sciadv.ads7187> (2025).

Acknowledgements We thank Y. Xiong for help with the FEP studies; and the UCSF Cryo-EM facility staff for training and technical assistance.

Author contributions B.K.S., X.X. and O.M. designed the project. X.X. and O.M. designed the analogues with help from N.D.L. α_2 Receptor studies were performed by X.X. and H.H., with guidance from P.G. α_2 -A-related behavioural analyses were conducted by J.M.B., with guidance from A.I.B. α_2 Structural studies were performed by K.S. under the supervision of A. Manglik. SERT uptake assays were carried out by X.-P.H. and J.W., with guidance from B.L.R. and ligand choice from Y.W. Mac1 biochemical assays were conducted by Y.U.D., M.G.V.S., M.E.D. and K.Z., and were supervised by A.A. Mac1 crystallography was performed by G.J.C. with guidance from J.S.F. CB2 receptor assays were performed by X.X. and C.I.-T., with guidance from A. Makriyannis and ligand advice from M.M.R. AmpC assays were conducted by X.X., F.L. and K.L.H. Y.S.M. coordinated and supervised small-molecule synthesis. All in vitro ADME and safety assays were supported by Y.H. and Y.K. through Bienta. FEP calculations and analyses were performed by D.S., guided by Y.Z. and R.A. O.M. performed statistical analyses for both ChEMBL and experimental data. X.X., O.M. and B.K.S. prepared the manuscript. B.K.S. supervised the project. All authors reviewed and approved the final manuscript.

Funding This work was supported by US NIH R35GM122481 (to B.K.S.), US DARPA ABC grant HR001123S0038 and US ARPA-H grant 1AY1AX000035 (principal investigator J.S.F.). O.M. was partially supported by US NIH postdoctoral fellowship F32GM154469. The UCSF cryo-EM equipment is partially supported by NIH grants S10OD020054, S10OD021741 and S10OD026881, and by the Howard Hughes Medical Institute.

Competing interests B.K.S. is co-founder of Epiodyne, BlueDolphin and Deep Apple Therapeutics; serves on the Scientific Advisory Board (SAB) for Schrödinger, Vilya Therapeutics and Frontier Discovery; and is on the Scientific Resource Board (SRB) of Genentech. B.L.R. is founder of Onsero Therapeutics. J.S.F. is a consultant to and a shareholder of Vilya Therapeutics and Relay Therapeutics. D.S., Y.Z. and R.A. are employed by Schrödinger Inc. Y.H., Y.K. and Y.S.M. are employed by Enamine Ltd. Y.S.M. is employed by Chemspace LLC. The remaining authors declare no competing interests.

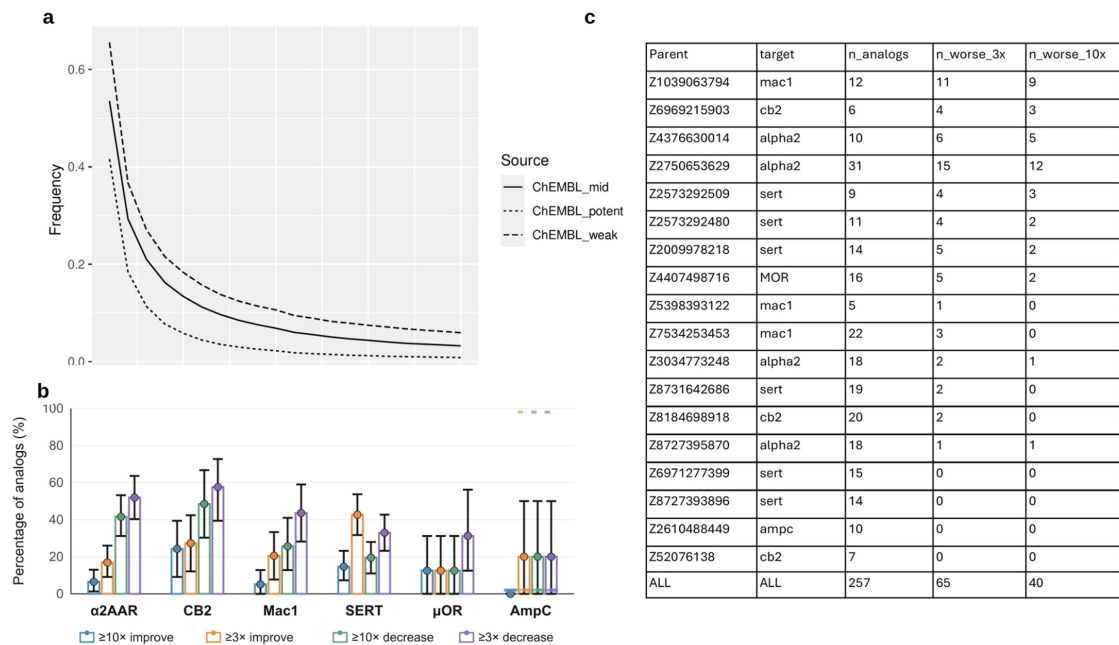
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-11013-5>.

Correspondence and requests for materials should be addressed to Bryan L. Roth, James S. Fraser or Brian K. Shoichet.

Peer review information Nature thanks Derek Lowe, Celine Valant and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

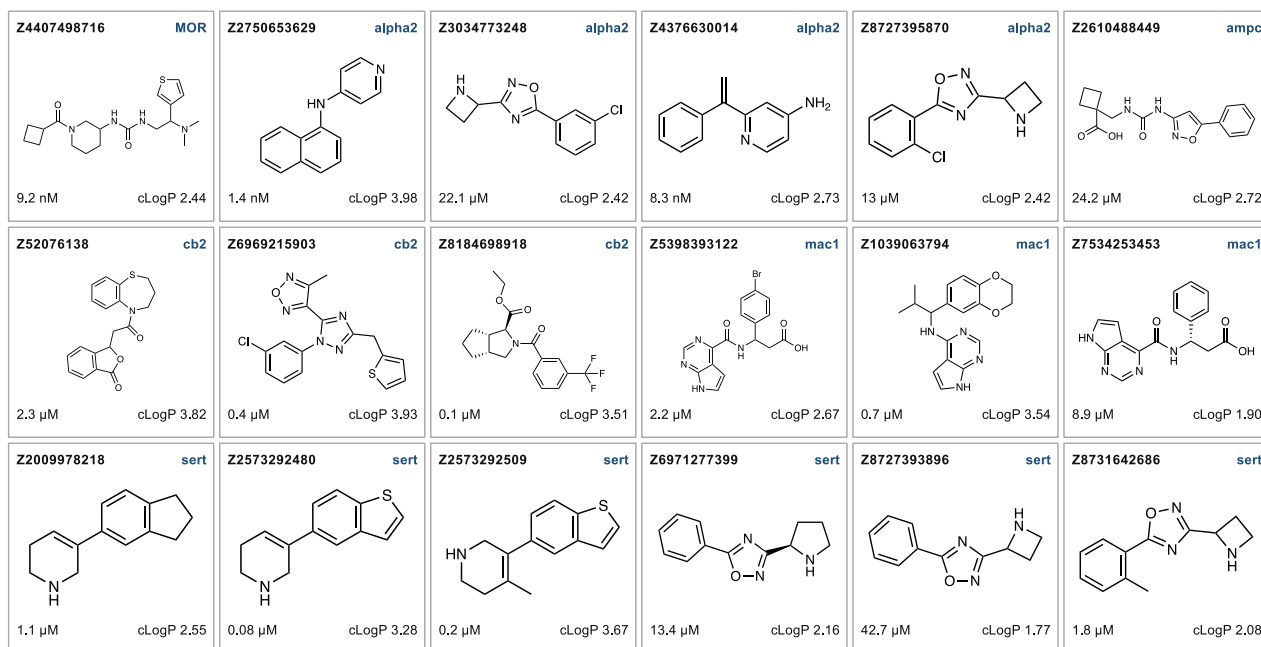
Reprints and permissions information is available at <http://www.nature.com/reprints>.



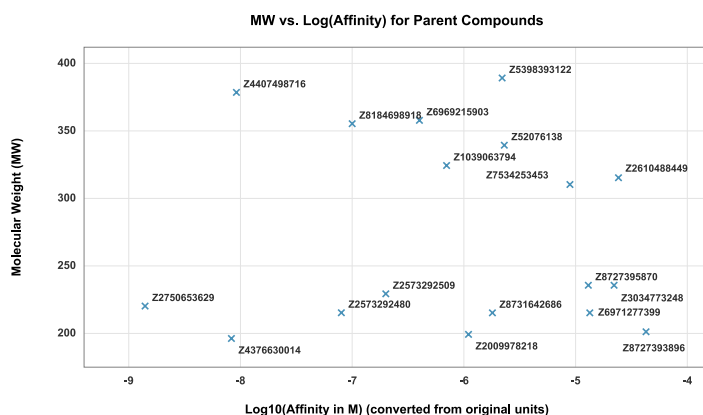
Extended Data Fig. 1 | Frequency of activity changes from single-atom substitutions in ChEMBL and in this study. a, Cumulative frequency of single non-hydrogen atom substitutions (CH₃, F, Cl, Br, N and OH) in ChEMBL that improve activity, stratified by that of the parent compound. “Potent” denotes parents with activity ≤32 nM, “mid” 32 nM–1 μM, and “weak” 1 μM–1 mM. b, For each target in this study, the percentage of analogs that improve or decrease activity by ≥3-fold or ≥10-fold versus their parent. Open bars and points show observed percentages; black error bars show two-sided 95% confidence

intervals estimated by percentile bootstrap (20,000 resamples). Sample sizes were n = 77, 33, 39, 82, 16 and 10 analogs for alpha2AAR, CB2, Mac1, SERT, μOR and AmpC, respectively. Individual binary outcomes are overlaid for AmpC (n = 10). Fold changes were rounded to the nearest integer before thresholding (for example, 9.6–9.9 counted as 10-fold). Compounds yielding no measurable curves were counted as ≥10-fold decreases. c, Parent-level summary of analog effects in this study (rounded fold-change).

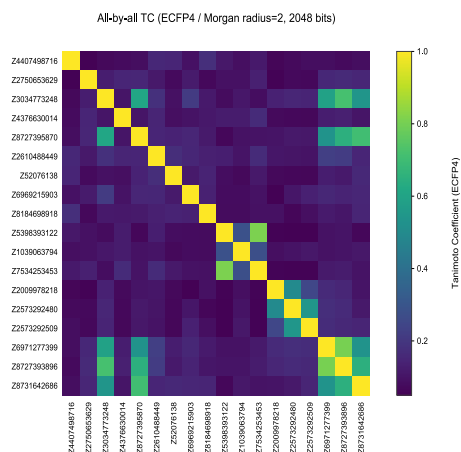
a



b

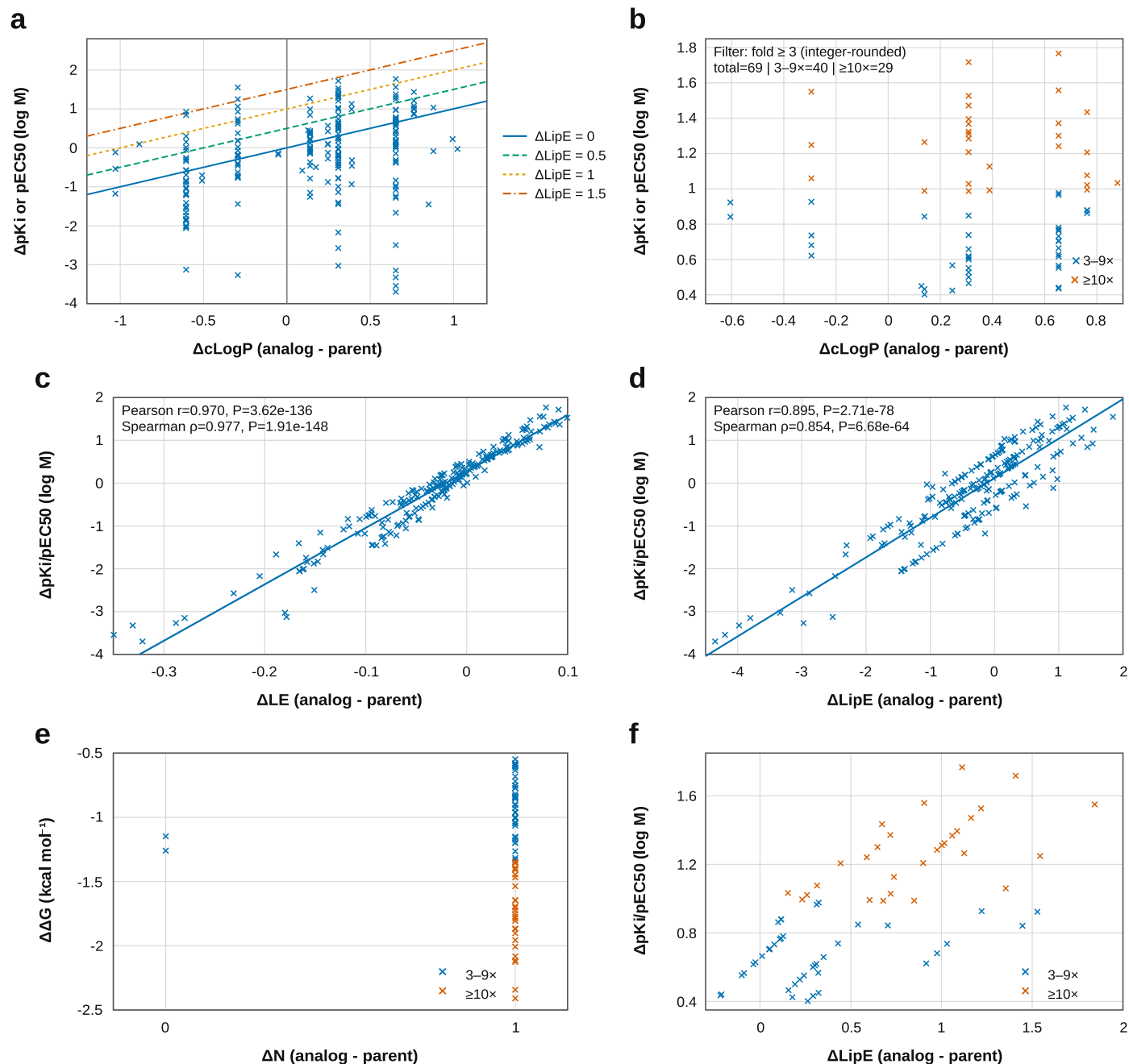


c



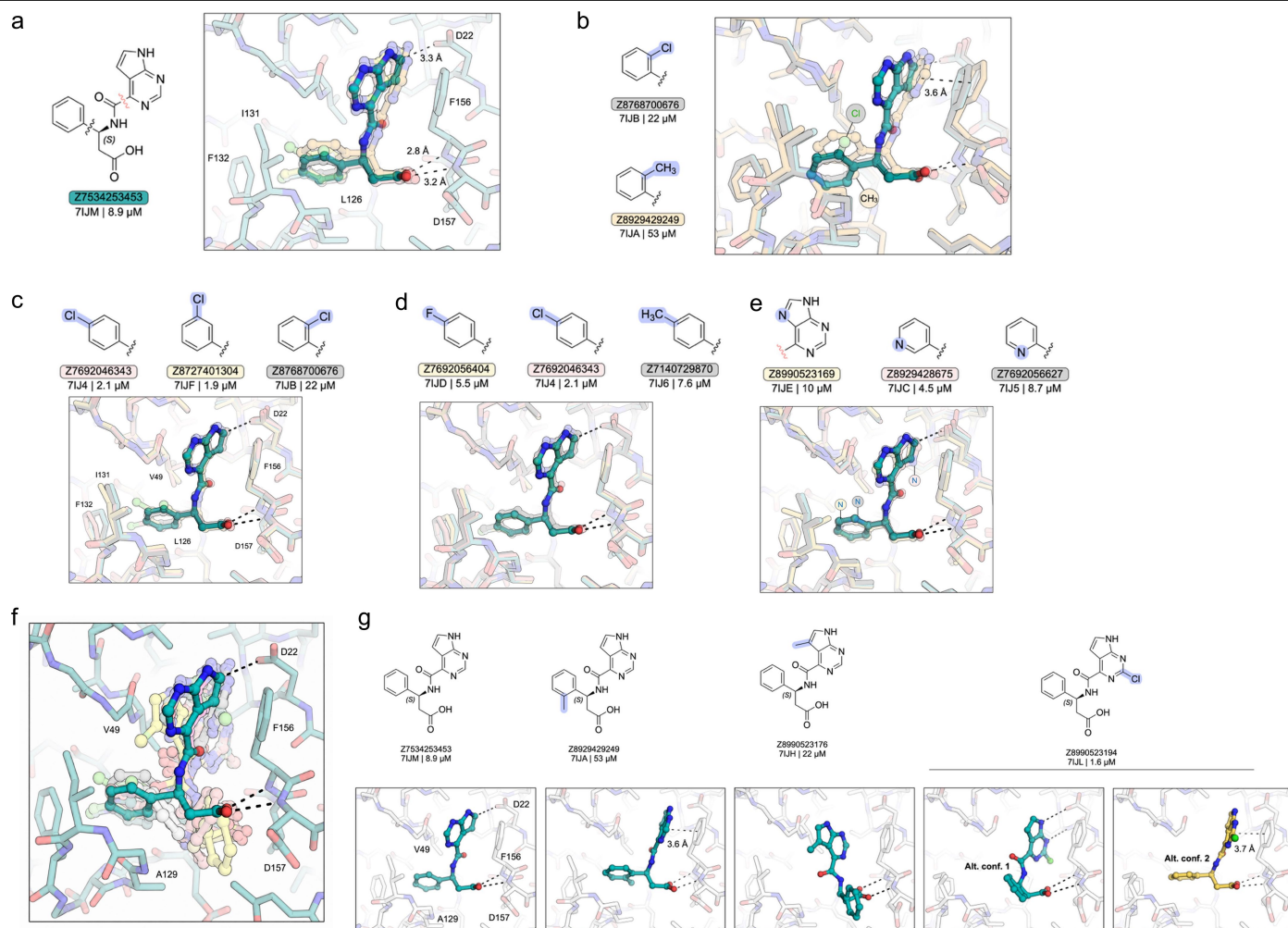
Extended Data Fig. 2 | Overview of Parent Compounds' Structures and Properties. a, 2D structures of the parent compounds, with their measured affinity or potency and calculated cLogP values. b, Molecular Weight (MW) and

Affinity or potency values for the parent compounds. c, Structural Similarity matrix showing the relationships between parent compounds based on molecular fingerprint comparisons.



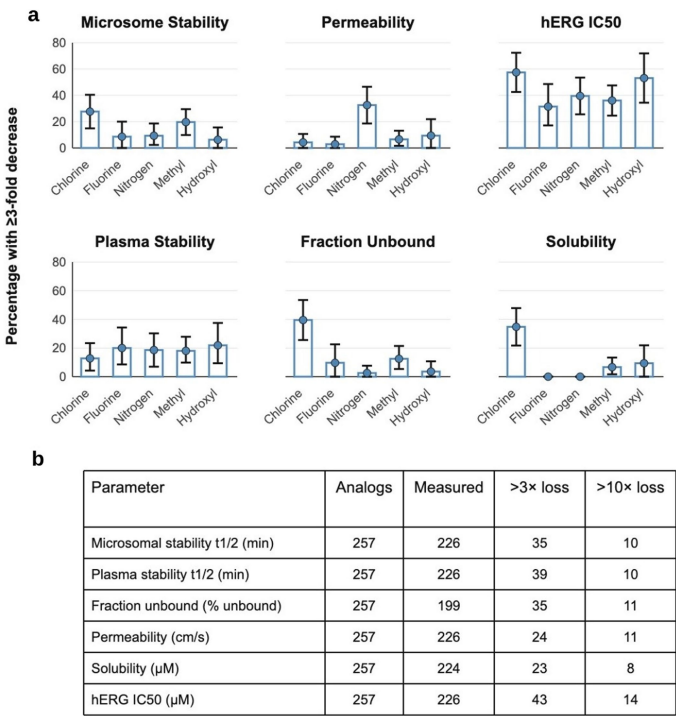
Extended Data Fig. 3 | Correlations between physicochemical descriptors and $\Delta pKi/pEC50$ (log M). a, Change in affinity or potency ($\Delta pActivity = pActivity(analog) - pActivity(parent)$, where $pActivity$ is pKi , $pIC50$ or $pEC50$) plotted against $\Delta cLogP$ (analog - parent) for all analog-parent pairs ($n = 220$). Diagonal isoclines indicate $\Delta LipE$ values of 0, 0.5, 1.0 and 1.5, where $LipE = pActivity - cLogP$; a point above a given isocline has a larger gain in lipophilic efficiency than the indicated value. b, ΔpKi or $pEC50$ (analog - parent) against $\Delta cLogP$ for all analog-parent pairs with $\geq 3\times$ rounded potency/affinity improvement; $\geq 10\times$ improvements are highlighted in red. c, Scatter plot of ΔpKi or $pEC50$ (log M) versus ΔLE (analog - parent), where $LE = 1.37 \times pKi$ or $pEC50/N_{heavy}$ ($n = 220$ analog-parent pairs). Pearson product-moment and Spearman rank correlations were assessed using two-sided tests ($R = 0.970$, $P = 3.62 \times 10^{-136}$; $\rho = 0.977$, $P = 1.91 \times 10^{-148}$). d, Scatter plot of ΔpKi or $pEC50$

(log M) versus $\Delta LipE$ (analog - parent), where $LipE = pKi$ or $pEC50 - cLogP$ ($n = 220$ analog-parent pairs). Pearson product-moment and Spearman rank correlations were assessed using two-sided tests ($R = 0.895$, $P = 2.71 \times 10^{-78}$; $\rho = 0.854$, $P = 6.68 \times 10^{-64}$). e, Binding free-energy change versus heavy-atom change for improved analogs. The y-axis shows $\Delta \Delta G$ (kcal/mol), computed from the affinity metric ($Ki/IC50/EC50$) as $\Delta G = RT \ln(K)$ (298 K) and $\Delta \Delta G = \Delta G_{analog} - \Delta G_{parent}$; negative $\Delta \Delta G$ indicates improved binding. The x-axis is ΔN (analog - parent; heavy-atom count difference). Points with rounded fold ≥ 10 are highlighted in red; rounded 3–9 are shown in blue. f, Affinity or potency gains versus changes in lipophilic efficiency. Scatter plot of ΔpKi or $pEC50$ (analog - parent) against $\Delta LipE$ (analog - parent) for all analog-parent pairs with $\geq 3\times$ rounded potency/affinity improvement; $\geq 10\times$ improvements are highlighted in red.



Extended Data Fig. 4 | Structural basis and variability of potency changes induced by single-atom modifications in Mac1 inhibitors. a, Chemical structure of the '3453 parent and overlay of X-ray crystal structures of 12 analogs. b, Alignment of '0676 (21 μ M) and '9249 (53 μ M) with the '3453 parent, showing the different binding poses of the ortho-chloro and ortho-methyl analogs. c, Comparison of '6343 (2.1 μ M), '1304 (1.9 μ M) and '0676 (21 μ M) shows that repositioning the aryl chloride disrupts packing with I131 and F132, reducing potency. d, Para-methyl ('9870, 7.6 μ M) and para-chloro ('6343, 2.1 μ M) substitutions have different effects despite similar sterics. The fluoro analog '6404 (5.5 μ M) binds more weakly than the chloro analog '6343,

consistent with weaker non-covalent interactions. e, Nitrogen substitutions (C $\rightarrow$ N) in '3169, '6627 and '8675 have different effects relative to their shared parent (8.9 μ M): '8675 (meta-N, 4.5 μ M) improves potency by approximately twofold, whereas '6627 (8.7 μ M) and '3169 (10 μ M) show little or no improvement. f, Overlay of the X-ray crystal structures of 12 ligand-bound complexes from the '3453 series, showing that single-atom modifications can cause substantial ligand movement or multiple binding poses. g, Examples of binding-pose shifts: '9249, '3176 and '3194 adopt markedly different poses relative to the '3453 parent, and two conformations are observed for '3194.



Extended Data Fig. 5 | Single-atom substitutions frequently worsen PK properties. a, Observed percentages of single-atom substitutions with ≥ 3 -fold decreases in PK properties. Open bars and points show observed percentages; black error bars show two-sided 95% confidence intervals estimated by percentile bootstrap (10,000 resamples). Exact numbers of measured analogs (Cl, F, N, CH₃ and OH, in that order) were: microsomal stability, n = 47, 35, 43, 61 and 32; permeability, n = 47, 35, 43, 61 and 32; hERG IC₅₀, n = 36, 28, 26, 52 and 17; plasma stability, n = 47, 35, 43, 61 and 32; fraction unbound, n = 43, 31, 39, 56 and 28; and solubility, n = 46, 35, 43, 60 and 32. b, Counts of analogs with measured PK parameters and >3 -fold or >10 -fold losses.

Extended Data Table 1 | Data collection and refinement statistics (molecular replacement)

	Z1039063794 (PDB 14AB)	Z5398393122 (PDB 7IIW)	Z7691912366 (PDB 7IIX)	Z8598075283 (PDB 7IIY)	Z8598075280 (PDB 7IIZ)	Z8598075284 (PDB 7IIJ)	Z8598075302 (PDB 7IIJ1)	Z8727396638 (PDB 7IIJ2)
Data collection								
Space group	P43	P43	P43	P43	P43	P43	P43	P43
Cell dimensions								
<i>a</i> , <i>b</i> , <i>c</i> (Å)	88.69, 88.69, 39.64	88.67, 88.67, 39.55	88.92, 88.92, 39.70	88.74, 88.74, 39.57	88.74, 88.74, 39.55	88.67, 88.67, 39.59	88.85, 88.85, 39.53	88.76, 88.76, 39.67
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)*	39.66-0.98 (1.02-0.98)	44.34-0.97 (1.01-0.97)	44.46-1.02 (1.06-1.02)	44.37-1.01 (1.05-1.01)	44.37-0.99 (1.03-0.99)	39.65-0.98 (1.02-0.98)	39.73-1.01 (1.05-1.01)	44.38-0.97 (1.01-0.97)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.03638 (1.162)	0.0564 (1.139)	0.06828 (2.139)	0.07219 (1.501)	0.04526 (0.9577)	0.04375 (1.513)	0.1217 (1.064)	0.05287 (1.076)
<i>I</i> / σ <i>I</i>	17.43 (0.88)	12.11 (0.73)	9.88 (0.61)	10.07 (0.83)	15.95 (1.32)	15.31 (0.81)	7.66 (1.13)	13.94 (0.84)
Completeness (%)	99.43 (94.88)	99.69 (97.35)	99.70 (97.77)	99.79 (99.02)	99.45 (96.15)	99.39 (94.54)	99.81 (98.70)	99.08 (91.50)
Redundancy	6.2 (3.8)	6.1 (3.6)	6.4 (5.9)	6.4 (5.5)	6.3 (4.8)	6.2 (4.3)	6.4 (5.5)	6.1 (3.7)
Refinement								
Resolution (Å)	0.98	0.97	1.02	1.01	0.99	0.98	1.01	0.97
No. reflections	175635	181108	157387	161035	170249	175303	161295	180870
<i>R</i> _{work} / <i>R</i> _{free}	0.1470 / 0.1642	0.1474 / 0.1666	0.1585 / 0.1770	0.1492 / 0.1702	0.1408 / 0.1577	0.1575 / 0.1784	0.1505 / 0.1653	0.1452 / 0.1641
No. atoms	3739	3757	3699	3805	3741	3734	3797	3872
Protein	3269	3274	3226	3308	3273	3238	3342	3397
Ligand/ion	44	36	39	39	35	39	37	49
Water	446	459	448	472	444	471	431	447
B-factors	18.44	18.65	19.28	19.7	17.46	18.58	18.93	16.07
Protein	16.82	16.92	17.58	17.94	15.84	16.86	17.43	14.7
Ligand/ion	14.29	16.42	16.86	16.6	15.27	14.6	14.46	10.43
Water	30.58	31.14	31.7	32.19	29.52	30.61	30.79	26.82
R.m.s. deviations								
Bond lengths (Å)	0.004	0.032	0.095	0.044	0.036	0.004	0.026	0.042
Bond angles (°)	0.76	1.13	2.26	1.66	1.19	0.76	1.28	1.95

*Values in parentheses are for the highest-resolution shell.

Extended Data Table 2 | Data collection and refinement statistics (molecular replacement)

	Z8727396582 (PDB 7UJ3)	Z7692046343 (PDB 7UJ4)	Z7692056627 (PDB 7UJ5)	Z7140729870 (PDB 7UJ6)	Z7691912473 (PDB 7UJ7)	Z8928826043 (PDB 7UJ8)	Z8928826033 (PDB 7UJ9)	Z8929429249 (PDB 7UJA)
Data collection								
Space group	P43	P43	P43	P43	P43	P43	P43	P43
Cell dimensions								
<i>a</i> , <i>b</i> , <i>c</i> (Å)	88.76, 88.76, 39.72	88.78, 88.78, 39.38	88.71, 88.71, 39.45	88.75, 88.75, 39.38	88.75, 88.75, 39.42	88.77, 88.77, 39.26	88.53, 88.53, 39.19	88.68, 88.68, 39.57
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)*	36.26-0.97 (1.01-0.97)	44.39-1.01 (1.05-1.01)	44.36-0.99 (1.03-0.99)	44.38-1.02 (1.06-1.02)	44.38-1.01 (1.05-1.01)	39.70-1.00 (1.04-1.00)	39.59-1.00 (1.04-1.00)	44.34-1.00 (1.04-1.00)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.05174 (1.234)	0.0572 (1.852)	0.03692 (1.117)	0.07428 (1.737)	0.05312 (0.7722)	0.063 (1.565)	0.06452 (1.431)	0.03096 (0.5122)
<i>I</i> / σ	13.43 (0.72)	11.96 (0.74)	18.44 (1.09)	8.69 (0.64)	14.20 (1.55)	10.80 (0.69)	12.39 (0.99)	22.79 (2.47)
Completeness (%)	99.54 (96.11)	99.65 (97.04)	99.18 (92.53)	99.68 (98.76)	99.80 (99.02)	99.54 (95.84)	99.46 (95.57)	99.67 (97.15)
Redundancy	6.0 (3.4)	6.3 (5.7)	6.3 (4.8)	6.3 (6.0)	6.4 (5.7)	6.3 (5.2)	6.3 (5.3)	6.4 (5.3)
Refinement								
Resolution (Å)	0.97	1.01	0.99	1.02	1.01	1	1	1
No. reflections	182045	160186	169237	155537	160480	164352	163046	165490
<i>R</i> _{work} / <i>R</i> _{free}	0.1469 / 0.1658	0.1537 / 0.1756	0.1472 / 0.1662	0.1553 / 0.1768	0.1432 / 0.1615	0.1578 / 0.1735	0.1585 / 0.1732	0.1434 / 0.1590
No. atoms	3731	3858	3766	3784	3826	4173	4244	3753
Protein	3236	3379	3302	3330	3379	3796	3870	3255
Ligand/ion	39	46	35	39	45	39	46	39
Water	471	451	441	430	420	353	346	474
B-factors	16.16	18.45	17.32	19.12	16.51	18.63	16.21	16.28
Protein	14.49	16.97	15.92	17.68	15.15	17.66	15.54	14.6
Ligand/ion	15.01	14.04	9.43	17.54	11.6	19.42	11.69	12.08
Water	27.7	29.79	28.2	30.39	27.77	28.98	24.08	27.99
R.m.s. deviations								
Bond lengths (Å)	0.004	0.005	0.026	0.005	0.005	0.005	0.004	0.004
Bond angles (°)	0.76	0.75	1.3	0.75	0.77	0.77	0.74	0.76

*Values in parentheses are for the highest-resolution shell.

Extended Data Table 3 | Data collection and refinement statistics (molecular replacement)

	Z8768700676 (PDB 7IJB)	Z8929428675 (PDB 7IJC)	Z7692056404 (PDB 7IJD)	Z8990523169 (PDB 7IJE)	Z8727401304 (PDB 7IJE)	Z8990527720 (PDB 7IJG)	Z8990523176 (PDB 7IJI)	Z8990523172 (PDB 7IJI)
Data collection								
Space group	P43	P43	P43	P43	P43	P43	P43	P43
Cell dimensions								
<i>a</i> , <i>b</i> , <i>c</i> (Å)	88.58, 88.58, 39.45	88.73, 88.73, 39.41	88.65, 88.65, 39.36	88.65, 88.65, 39.53	88.77, 88.77, 39.40	88.57, 88.57, 39.48	88.86, 88.86, 39.49	88.88, 88.88, 39.51
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)*	39.62-1.00	39.69-1.00	39.64-1.00	44.33-1.00	33.37-1.00	44.29-1.02	44.43-1.00	44.44-1.02
<i>R</i> _{sym} or <i>R</i> _{merge}	(1.04-1.00)	(1.04-1.00)	(1.04-1.00)	(1.04-1.00)	(1.04-1.00)	(1.06-1.02)	(1.04-1.00)	(1.06-1.02)
<i>I</i> / σ	0.03626 (0.5552)	0.03932 (1.702)	0.03193 (1.025)	0.04505 (1.224)	0.03926 (1.832)	0.03189 (0.3354)	0.04984 (1.279)	0.04055 (1.789)
Completeness (%)	20.58 (2.27)	16.41 (0.80)	21.22 (1.36)	15.52 (1.04)	16.60 (0.74)	25.43 (4.31)	13.60 (0.95)	15.94 (0.90)
Redundancy	99.74 (97.57)	99.56 (96.24)	99.65 (96.89)	99.68 (97.13)	99.39 (94.64)	99.86 (99.51)	99.51 (95.76)	99.74 (98.53)
	6.4 (5.3)	6.4 (5.4)	6.4 (5.3)	6.3 (5.3)	6.4 (5.3)	6.5 (6.2)	6.3 (5.3)	6.5 (6.2)
Refinement								
Resolution (Å)	1	1	1	1	1	1.02	1	1.02
No. reflections	164729	164766	164447	165231	164639	155509	165540	156606
<i>R</i> _{work} / <i>R</i> _{free}	0.1381 / 0.1573	0.1598 / 0.1817	0.1451 / 0.1648	0.1448 / 0.1647	0.1557 / 0.1766	0.1384 / 0.1530	0.1464 / 0.1638	0.1615 / 0.1816
No. atoms	3860	3817	3816	3831	3817	3901	3711	3857
Protein	3359	3356	3330	3336	3358	3408	3222	3404
Ligand/ion	46	45	46	45	46	46	39	39
Water	473	434	458	468	431	465	465	429
B-factors	15.66	19.72	17.47	17.1	19.06	14.36	18.17	19.93
Protein	14.2	18.34	16.02	15.62	17.74	13.07	16.46	18.47
Ligand/ion	9.95	12.31	11.94	11.86	14.89	9.19	12.41	14.62
Water	26.4	30.91	28.33	27.95	29.66	24.19	30.33	31.84
R.m.s. deviations								
Bond lengths (Å)	0.004	0.025	0.004	0.005	0.005	0.004	0.005	0.092
Bond angles (°)	0.77	1.3	0.76	0.78	0.76	0.78	0.77	2.17

*Values in parentheses are for the highest-resolution shell.

Extended Data Table 4 | Data collection and refinement statistics (molecular replacement)

	Z8990919037 (PDB 7IJJ)	Z8990919048 (PDB 7IJK)	Z8990523194 (PDB 7IJL)	Z8207599372 (PDB 14AM)	Z8756044539 (PDB 14AN)	Z9187344358 (PDB 14AO)	Z9187344338 (PDB 14AP)	Z7534253453 (PDB 7IJM)
Data collection								
Space group	P43	P43	P43	P 43	P 43	P 43	P 43	P43
Cell dimensions								
<i>a</i> , <i>b</i> , <i>c</i> (Å)	88.71, 88.71, 39.46	88.69, 88.69, 39.40	88.81, 88.81, 39.64	88.71, 88.71, 39.45	88.75, 88.75, 39.43	88.77, 88.77, 39.63	88.67, 88.67, 39.30	88.76, 88.76, 39.38
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)*	36.05-0.98	39.66-0.98	39.71-1.01	36.05-0.99	44.38-0.98	39.70-1.02	39.65-1.00	39.70-1.01
<i>R</i> _{sym} or <i>R</i> _{merge}	(1.02-0.98)	(1.02-0.98)	(1.05-1.01)	(1.03-0.99)	(1.02-0.98)	(1.06-1.02)	(1.04-1.00)	(1.05-1.01)
<i>I</i> / σ	0.04019 (1.282)	0.03954 (0.6102)	0.04512 (2.454)	0.03622 (1.34)	0.03458 (1.691)	0.05749 (2.14)	0.04611 (1.939)	0.05101 (1.889)
Completeness (%)	16.27 (0.82)	17.80 (1.55)	14.42 (0.61)	18.05 (0.88)	16.67 (0.64)	12.76 (0.76)	13.81 (0.73)	14.06 (0.77)
Redundancy	99.53 (95.82)	99.58 (96.36)	99.55 (96.58)	99.88 (99.12)	98.99 (90.67)	99.75 (98.33)	99.66 (97.37)	99.86 (98.86)
	6.2 (4.0)	6.2 (4.0)	6.4 (5.8)	6.3 (4.5)	6.2 (4.1)	6.4 (6.1)	6.4 (5.2)	6.4 (5.8)
Refinement								
Resolution (Å)	0.98	0.98	1.01	0.99	0.98	1.02	1	1.01
No. reflections	175117	174854	161129	170458	174211	156628	164244	160473
<i>R</i> _{work} / <i>R</i> _{free}	0.1524 / 0.1723	0.1380 / 0.1531	0.1635 / 0.1841	0.1471 / 0.1646	0.1664 / 0.1836	0.1597 / 0.1813	0.1737 / 0.1925	0.1582 / 0.1790
No. atoms	3866	3732	3732	3644	3807	3812	3804	3895
Protein	3390	3204	3276	3190	3367	3326	3390	3437
Ligand/ion	47	37	36	44	47	63	43	46
Water	448	504	432	429	415	454	390	431
B-factors								
Protein	18.12	16.02	20.69	18.44	20.79	17.65	20.56	17.65
Ligand/ion	16.71	14.33	19.12	16.83	19.41	16.16	19.43	16.46
Water	12.27	7.21	18.21	16.65	17.41	12.72	14.97	12.43
R.m.s. deviations	29.14	27.2	32.71	30.56	32.22	28.89	30.74	27.41
Bond lengths (Å)	0.025	0.004	0.026	0.004	0.004	0.005	0.098	0.004
Bond angles (°)	1.3	0.77	1.32	0.74	0.73	0.75	1.69	0.75

*Values in parentheses are for the highest-resolution shell.

Extended Data Table 5 | Cryo-EM data collection, refinement and validation statistics

	4905-bound α 2AAR complex with G proteins and scFv16 (EMD-71719) (PDB 9PLO)	4905-bound locally refined α 2AAR (EMD-71718) (PDB 9PLN)
Data collection and processing		
Microscope and detector		
Microscope	Thermo Scientific Krios G3	Thermo Scientific Krios G3
Detector	Gatan K3 with BioQuantum filter	Gatan K3 with BioQuantum filter
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure ($e^-/\text{\AA}^2$)	47.7	47.7
Defocus range (μm)	-0.8 to -2.2	-0.8 to -2.2
Pixel size ($\text{\AA}$)	0.8189	0.8189
Micrographs (no.)	10,305	10,305
Symmetry imposed	C1	C1
Initial particle images (no.)	6,609,807	6,609,807
Final particle images (no.)	323,827	323,827
Map resolution ($\text{\AA}$)	2.75	2.8
FSC threshold	0.143	0.143
Map resolution range ($\text{\AA}$)	—	—
Refinement		
Initial model used (PDB code)	7EJ8	7EJ8
Model resolution ($\text{\AA}$)	2.75	2.7
FSC threshold	0.5	0.5
Model resolution range ($\text{\AA}$)	—	—
Map sharpening B factor ($\text{\AA}^2$)	-110.6	-115.5
Model composition		
Chains	5	1
Non-hydrogen atoms	8,635	2,117
Protein residues	1,109	266
Ligands	1	1
B factors ($\text{\AA}^2$)		
Protein (min/max/mean)	78.15/211.79/128.69	79.96/157.60/113.70
Ligand (min/max/mean)	119.15/137.53/129.82	96.23/109.98/103.57
R.m.s. deviations		
Bond lengths ($\text{\AA}$)	0.005	0.006
Bond angles ($^\circ$)	1.104	1.214
Validation		
MolProbity score	1.35	1.21
Clashscore	3.22	4.27
Poor rotamers (%)	0.43	0.90
Ramachandran plot		
Favored (%)	96.43	98.08
Allowed (%)	3.57	1.92
Disallowed (%)	0.00	0.00

Dashes indicate values that were not reported.

Reporting Summary

Nature Portfolio 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 Portfolio 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ 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 ALS beamline 8.3.1: Blu-Ice 5.

Data analysis Confirmed software and version details are listed on appended page 6.

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All primary data are available through the manuscript, Extended Data, Supplementary Information and Source Data. Cryo-EM structures/maps: PDB 9PLO/EMD-71719 and PDB 9PLN/EMD-71718. Mac1 structures and structure factors: PDB 14AB, 7IIW, 7IIX, 7IY, 7IIZ, 7IJ0, 7IJ1, 7IJ2, 7IJ3, 7IJ4, 7IJ5, 7IJ6, 7IJ7, 7IJ8, 7IJ9, 7IJA, 7IJB, 7IJC, 7IJD, 7IJE, 7IJF, 7IJG, 7IJH, 7IJI, 7IJJ, 7IJK, 7IJL, 14AM, 14AN, 14AO, 14AP and 7IJM. ChEMBL database version 34 is available at <https://doi.org/10.6019/CHEMBL.database.34>. Compounds and catalogue identifiers are listed in Supplementary Table 1.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

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](https://www.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 synthesized 257 analogs of 18 parent compounds. Sample sizes were not predetermined by statistical methods but were driven by compound availability and feasibility. For mouse studies, no formal a priori power calculation was performed; group sizes followed prior assay experience and literature precedent.
Data exclusions	No data were excluded.
Replication	Fig. 5b comprises three independent experiments, each with two technical replicates. Mouse behavioral experiments were performed in two independent cohorts.
Randomization	Mice were randomly assigned to treatment and control groups as described in the Methods.
Blinding	For mouse experiments, behavioral assessment was performed by an experimenter different from the injector; behavioral investigators were therefore blinded to treatment.

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
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#) and [Sex and Gender in Research](#)

Cell line source(s)

Sf9: Expression Systems 94-001S; Hi5: Expression Systems 94-002S; HEK293: ATCC CRL-1573; HEK293T: ATCC CRL-3216.

Authentication

All cell lines were authenticated by their respective suppliers.

Mycoplasma contamination

The cell lines were not tested for mycoplasma contamination by the authors.

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

no commonly misidentified cell lines were used.

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.

Animals and other organisms

Policy information about studies involving animals and other organisms

Laboratory animals	Male C57BL/6J mice, 7-8 weeks old, obtained from The Jackson Laboratory (JAX strain #000664).
Housing conditions	Five mice per cage; standard 12:12-hour light-dark cycle; 20.7 C (69.3 F); 58% humidity.
Ethics oversight	University of California, San Francisco Institutional Animal Care and Use Committee (IACUC), protocol #AN208219.
Randomization	Animals were randomly assigned to treatment and control groups.
Blinding	The experimenter assessing behavior was different from the experimenter administering injections and was blinded to treatment.
Replication	Behavioral experiments were conducted in two independent cohorts.
Wild animals	Not applicable.
Field-collected samples	Not applicable.

The corresponding information has also been added to the manuscript Methods and Fig. 5 legend.

Antibodies

Policy information about antibodies used in the study

Antibody/reagent	Anti-His6-Eu3+ cryptate, used as the HTRF donor reagent.
Supplier	PerkinElmer.
Catalogue number	AD0402.
Application	Mac1 HTRF displacement assay for detection of His6-tagged protein.
Dilution	1:20,000.
Validation	Validated by the manufacturer for HTRF detection of His6-tagged proteins.

Supplier, catalogue number, application, dilution and manufacturer validation are also stated in the author response.

Software and code details

Confirmed software and instrument-software versions

X-ray data collection	Blu-Ice 5 at ALS beamline 8.3.1.
X-ray processing	XDS, dated 19 Jan 2025, BUILT=20250430; Aimless 0.8.2.
Density/model analysis	PanDDA 0.2.14; COOT 0.8.9.2; phenix.refine 1.21.2_5419.
Structural visualization	PyMOL 3.1.8; ChimeraX 1.8.
Cryo-EM processing	cryoSPARC 4.5.3.
Mac1 instruments	Echo 650 Liquid Handler 3.1.0; EnVision 2105-0010 software 1.14.3049.1193.
Plate-reader analysis	SoftMax Pro 7.
Statistical/data analysis	GraphPad Prism 9.0, 10.0 and 10.0.2; Microsoft Excel 16.99.2.
FEP calculations	Schrodinger FEP+ 2025-2 with the OPLS4 force field.
Cheminformatics/analysis	Corina 3.60; Omega 3.0.1; ChemAxon JChem 21.13; RDKit 2020.09.1; Python 2.7 and 3.7; R 4.2.0 (ggplot2).