

## Low intrinsic efficacy for G protein activation can explain the improved side-effect profile of new opioid agonists

**Authors:** Alexander Gillis <sup>1§</sup>, Arisbel B Gondin <sup>2§</sup>, Andrea Kliewer <sup>3</sup>, Julie Sanchez <sup>2,9,10</sup>, Herman D Lim <sup>2</sup>, Claudia Alamein <sup>1</sup>, Pradeep Manandhar <sup>4</sup>, Marina Santiago <sup>4</sup>, Sebastian Fritzwanker <sup>3</sup>, Frank Schmidel <sup>3</sup>, Timothy A Katte <sup>5</sup>, Tristan Reekie <sup>5</sup>, Natasha L Grimsey <sup>6</sup>, Michael Kassiou <sup>5</sup>, Barrie Kellam <sup>7</sup>, Cornelius Krasel <sup>8</sup>, Michelle L Halls <sup>2</sup>, Mark Connor <sup>4</sup>, J Robert Lane <sup>9,10</sup>, Stefan Schulz <sup>3</sup>, Macdonald J Christie <sup>1\*</sup>, Meritxell Canals <sup>2,9,10\*</sup>

### Affiliations:

<sup>1</sup> Discipline of Pharmacology, School of Medical Sciences, University of Sydney, NSW 2006, Australia.

<sup>2</sup> Drug Discovery Biology Theme, Monash Institute of Pharmaceutical Sciences, Monash University, Victoria 3052, Australia.

<sup>3</sup> Institute of Pharmacology and Toxicology, Jena University Hospital, Friedrich-Schiller-University, 07747 Jena, Germany.

<sup>4</sup> Department of Biomedical Sciences, Macquarie University, NSW, 2109, Australia.

<sup>5</sup> School of Chemistry, The University of Sydney, NSW, 2006, Australia.

<sup>6</sup> Department of Pharmacology and Clinical Pharmacology, School of Medical Sciences, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

<sup>7</sup> School of Pharmacy, Centre for Biomolecular Sciences, University of Nottingham, Nottingham NG7 2RD, United Kingdom

<sup>8</sup> Department of Pharmacology and Toxicology, Philipps-University Marburg, D-35043 Marburg, Germany.

<sup>9</sup> Division of Physiology, Pharmacology and Neuroscience, School of Life Sciences, Queen's Medical Centre, University of Nottingham, Nottingham, United Kingdom.

<sup>10</sup> Centre of Membrane Proteins and Receptors, Universities of Birmingham and Nottingham, The Midlands, United Kingdom.

<sup>§</sup> These authors contributed equally to this work.

\*Corresponding authors. Email: [m.canals@nottingham.ac.uk](mailto:m.canals@nottingham.ac.uk) (MC) and [mac.christie@sydney.edu.au](mailto:mac.christie@sydney.edu.au) (MJC)

**One sentence summary:** Low intrinsic efficacy can explain the reduced side effects of [apparently](#) biased  $\mu$ -opioid receptor agonists.

### Editor's Summary:

The antinociceptive effects of the  $\mu$ -opioid receptor (MOR) have made this receptor a powerful target for pain management, but serious side effects limit their use. Because the MOR is a G protein-coupled receptor (GPCR) that stimulates intracellular signaling through both G proteins and  $\beta$ -arrestins, biased agonists that primarily stimulate G protein signaling have been developed to promote pain relief without stimulating unwanted side effects. Gillis *et al.* compared the biochemical, signaling, and physiological properties of some of these G protein-biased MOR agonists with those of opioids that do not exhibit biased signaling. The authors found that the observed reductions in side effects could be explained by the low intrinsic efficacy of the biased agonists rather than their signaling bias per se. These findings suggest strategies for developing new agonists that relieve pain with fewer unwanted side effects.

**Commented [AV1]:** This is my one-sentence summary of the paper. Feel free to suggest revisions, but keep in mind that this sentence is limited to 120 characters or fewer, including spaces.

**Commented [AV2]:** This is my summary of the paper. Feel free to suggest revisions, but keep in mind that this paragraph is limited to 150 words or fewer (preferably much fewer).

**Commented [M3R2]:** I am happy with it

**Commented [MC4R2]:** I am happy with it as well

## Abstract

Biased agonism at G protein-coupled receptors (GPCRs) describes the phenomenon whereby some drugs can activate some downstream signaling activities to the relative exclusion of others. Descriptions of biased agonism focusing on the differential engagement of G proteins versus  $\beta$ -arrestins are commonly limited by the small response windows obtained in pathways that are not amplified or are less effectively coupled to receptor engagement, such as  $\beta$ -arrestin recruitment. At the  $\mu$ -opioid receptor (MOR), G protein-biased ligands have been proposed to induce less constipation and respiratory depressant side effects than opioids commonly used to treat pain. However, it is unclear whether these improved safety profiles are due to a reduction in  $\beta$ -arrestin-mediated signaling or, alternatively, to their low intrinsic efficacy in all signaling pathways. Here, we systematically evaluated the most recent and promising MOR biased ligands and assessed their pharmacological profile against existing opioid analgesics in assays not confounded by limited signal windows. We found that oliceridine, PZM21, and SR-17018 had low intrinsic efficacy. We also demonstrated a strong correlation between measures of efficacy for receptor activation, G protein coupling and  $\beta$ -arrestin recruitment for all tested ligands. By measuring the ~~anti-noiceptive~~ [antinociceptive](#) and respiratory depressant effects of these ligands, we showed that low intrinsic efficacy of opioid ligands can explain an improved side effect profile. Our results establish a possible alternative mechanism underlying the improved therapeutic windows described for new opioid ligands, which should be taken into account for future descriptions of ligand action at this important therapeutic target.

**Commented [AV5]:** I have not edited the Abstract yet – I typically do this section last.

**Commented [MC6]:** I am happy with this phrasing but one reviewer had concerns about this. J. Foley agreed the best sentence would be “our results suggest an alternative mechanism”.

## Introduction

Agonists of the  $\mu$ -opioid receptor (MOR), such as morphine and the synthetic opioid fentanyl, are mainstay analgesics for the treatment for severe acute pain. Unfortunately, MOR agonists also elicit several on-target adverse effects that severely limit their use, including respiratory depression, constipation, and the development of tolerance and addiction. The MOR is predominantly coupled to the  $G_{i/o}$  protein family, which signals by inhibiting the production of cyclic AMP (cAMP) by adenylyl cyclase (AC) [via  \$G\alpha\$ -subunits](#), and activating G protein-coupled inwardly rectifying potassium (GIRK) channels [through  \$G\beta\gamma\$  subunits](#), among other effectors (1) [\[1\]](#). Similar to most other G protein-coupled receptors (GPCRs), MOR signalling is regulated by the phosphorylation of intracellular C-terminal serine and threonine residues, which stabilizes the binding of  $\beta$ -arrestins. This family of cytosolic scaffolding proteins are understood to terminate G protein signalling and mediate the formation of endocytic complexes and receptor internalization.

$\beta$ -arrestin 2 has attracted much interest in regard to its interactions with MOR due to its suggested involvement in opioid side effects. Early studies showed enhanced morphine analgesia in  $\beta$ -arrestin 2 knockout mice with greatly diminished respiratory depression and constipation (2, 3). This has led to the hypothesis that a putative arrestin-dependent mechanism downstream of the MOR mediates the unwanted side effects of opioids (4). However, the nature of this proposed signal altering respiratory and gastrointestinal function has not been demonstrated. This 'arrestin hypothesis' has been challenged by the persistence of morphine- and fentanyl-induced side effects, including respiratory depression and constipation, in a knock-in mouse expressing a phospho-deficient MOR mutant that is unable to recruit  $\beta$ -arrestin (5) as well as by the persistence of morphine-induced respiratory depression in the  $\beta$ -arrestin 2 knockout mouse (6) [\[6\]](#) [\(Kliwer et al. Br J Pharmacol In press\)](#). Additionally, there is robust physiological evidence demonstrating that classical G protein signaling from the MOR contributes substantially to respiratory depression (7, 8), as well as to other side effects, such as constipation (9, 10).

The proposed role of  $\beta$ -arrestins in the unwanted effects of MOR agonists has led to the development of MOR ligands that do not recruit  $\beta$ -arrestin 2 to the receptor, under the assumption that this would avoid on-target side effects mediated by the proposed  $\beta$ -arrestin-dependent mechanism. Biased or functionally selective agonists engage a subset of signaling pathways to the exclusion of others, as in the case of G protein-biased ligands that maintain the classical G protein signal while reducing interactions with  $\beta$ -arrestin 2. The prototypical such ligand of the MOR is oliceridine (TRV130), initially reported to have an improved pre-clinical profile over morphine in rodent studies (11) and now in clinical trials. Further studies have reinforced that oliceridine induces constipation and abuse-related effects in rodents (12, 13), and although it still induces respiratory depression and constipation in humans, these side effects may be reduced when compared to morphine, giving a potentially wider therapeutic window (14, 15). Another MOR ligand proposed to be G protein-biased, PZM21, was discovered using in silico docking to the inactive MOR crystal structure (16). Studies in cell lines as well as in neurons of MOR-Venus knock-in mice have further shown that both oliceridine and PZM21 induce very limited  $\beta$ -arrestin recruitment or receptor internalization (17). The authors suggested that oliceridine and PZM21 have a signaling signature similar to buprenorphine, an extremely low-efficacy MOR agonist with reduced risk

Formatted: Font: Symbol

Commented [AV7]: Isn't this through  $G\beta\gamma$  and not through  $G\alpha$ ?

Formatted: Highlight

Commented [AV8]: Incorporate this to the reference list and cite it by number. You will need to reorder the references to insert this one. Remember that the references should be listed in the Reference List in the order in which they are cited in the paper and called out in the text by number.

of severe adverse effects. Although initially reported not to induce respiratory depression (16)(14), a later study found that PZM21 did slow respiratory frequency in mice to a similar extent as did morphine (18). Schmid *et al.* (19) generated a series of MOR ligands with increasing degrees of separation between G protein coupling, as measured by GTP $\gamma$ S binding and inhibition of cAMP accumulation, and  $\beta$ -arrestin coupling, and correlated the quantitative bias factors to the therapeutic window of these compounds. Of these ligands, SR-17018 was the compound with the most extreme apparent bias and the best separation between anti-nociception and respiratory side effects, because it was reported to not alter blood oxygen saturation or respiratory frequency in mice (19).

New MOR agonists proposed to be G protein-biased appear to have other pharmacological properties that could potentially be related to their safety profiles (20), and the mechanism through which reduced recruitment of  $\beta$ -arrestin 2 to the MOR could reduce side effects remains unclear. Characterization of biased agonists is complicated by the difficulty in accurately quantifying agonist activity for a given signaling pathway. The *de facto* standard for bias quantification is the operational model of agonism (21). This model, by design, mathematically accounts for the distinct pharmacological parameters of affinity (ligand binding to a target) and efficacy (ligand-induced activation of that target). Operational model analysis of bias is routinely performed by estimating a combined “transducer coefficient” as an index of intrinsic relative activity of an agonist for a given pathway. This can be normalized to a reference agonist to allow comparison of activity between pathways, through quantitative test agonist bias factors, presumably without system bias. However, there are important confounding factors that must be considered in the application of operational analysis to any system. Ligand binding kinetics and the temporal pattern of signaling processes are not accounted for in this analysis and can have a profound influence on apparent bias (22). Additionally, test compounds with very low activity for a particular end-point severely confound accurate quantification. Assays in which agonists produce no or minimal signal prevent satisfactory curve fit and, therefore, any robust estimation of bias using the operational model.

Regardless of the model, identification of biased agonists requires estimation of agonist efficacy and affinity independent of system and observational bias, such as cellular background and assay conditions. Initial descriptions of proposed G protein-biased MOR agonists did not specifically quantify agonist efficacy, giving either combined transduction coefficients (16, 19) or not directly quantifying bias (11). Assays of G protein activity are commonly confounded by the presence of receptor reserve (also known as spare receptors), which is the condition whereby an agonist needs to activate only a small fraction of the existing receptor population to produce the maximal system response (23, 24)(21, 24). In the presence of high receptor reserve, most test agonists reach a similar maximal response, preventing straightforward determination of relative efficacy (23). It is only when there is low receptor reserve that the differences in efficacy become apparent. In such cases, irreversible receptor inactivation removes receptor reserve and allows the quantification of ligand-intrinsic efficacy (23). Similarly, low-efficacy agonists, when compared between G protein assays, which measure highly amplified signalling, and  $\beta$ -arrestin recruitment measurements, which measure unamplified signals, may appear biased due to assay conditions (20). As such, recent studies showing the low efficacy of putatively biased agonists (18, 25, 26) should prompt a re-examination of the signaling profile of these biased MOR agonists, given that this factor was not initially identified.

**Commented [MC9]:** I could add another reference to the definition if required. Neubig et al 2003 Pharmacol Reviews 55:597-606

**Commented [AV10R9]:** If you feel that a second reference is necessary, then please feel free to insert it into the reference list in the appropriate place and refer to it here by that number. Remember that the references should be listed in the Reference List in the order in which they are cited in the paper and called out in the text by number.

**Commented [M11R9]:** I think the second reference should be added and have included it in the list. I cannot touch the numbering because the version is still linked to endnote. Meri, can you fix the numbering in the endnote file. I hate endnote

Similar to bias factors, accurate quantification of a pre-clinical ‘therapeutic window’ of lead compounds in vivo is critical to drug development. In this context, therapeutic window indicates the separation between the dose of a compound producing analgesic effect and the dose resulting in side effects such as respiratory depression or constipation. Therapeutic window is typically quantified pre-clinically by comparison of the potency of a compound for each response. The activity of test agonists in animal assays of anti-nociception and side effects is highly context-dependent. As such, measures of agonist bias and safety profiles using different in vitro assays, varying in vivo models, and inconsistent mathematical analyses does not allow for direct comparison of biased MOR agonists described in separate studies.

Here, we provide a systematic evaluation of the most recent and promising MOR biased ligands and assess their pharmacological profiles against existing opioid analgesics in assays that are not confounded by limited signal windows. We found that oliceridine, PZM21, and SR-17018 have very low intrinsic efficacy. We also demonstrated a strong correlation between measures of efficacy for receptor activation, G protein coupling, and  $\beta$ -arrestin recruitment for all tested ligands, including those that have previously been described as biased. Finally, by measuring the [antinociceptive](#) and respiratory depressant effects of these ligands, we showed that low intrinsic efficacy of the so-called biased ligands can explain their improved side effect profiles in terms of a continuum of existing analgesics. Our results therefore suggest an alternative mechanism underlying the improved therapeutic windows described for novel opioid ligands, which should be taken into account for future descriptions of ligand action at this important therapeutic target.

## Results

### *In vitro responses of different opioid ligands reveal a spectrum of maximal effects*

To understand the relationship between the intrinsic efficacy of opioids for different signalling pathways and, eventually, to assess their correlation with therapeutic indices, we constructed concentration-response curves for several opioid ligands and multiple downstream signaling pathways [proximally](#) linked to MOR activation in HEK293 cells (Fig. 1A). An analogue of the endogenous opioid peptide Met-enkephalin, DAMGO, was used as a reference agonist (16, 19, 24, 27). Existing clinical opioid agonists fentanyl, methadone, morphine, oxycodone and buprenorphine were profiled along with the three most recently described G protein-biased agonists oliceridine, PZM21, and SR-17018 (11, 16, 19). The ability of ligands to induce the active conformation of MOR was monitored through bioluminescence resonance energy transfer (BRET) assays for the recruitment of a conformationally selective nanobody (Nanobody 33-Venus, Nb33-Venus) to the receptor (Fig. 1B) (28). Similarly, coupling to and activation of  $G_{\alpha i}$ -proteins was assessed using, respectively, BRET assays with truncated, soluble ‘mini’  $G_i$ -proteins fused to a Venus fluorescent protein (m $G_{\alpha i}$ -Venus, Fig. 1C) (29) or previously described  $G_{\alpha i 2}$  activation BRET-based biosensors (Fig. 1D) (30, 31). Subsequent  $G_{\alpha i}$ -mediated inhibition of forskolin-induced cAMP production was monitored using the cAMP-YFP-Epac-RLuc (CAMYEL) BRET-based sensor (Fig. 1E) (32).  $G\beta\gamma$ -mediated activation of GIRK channels in response to MOR ligands was measured using a membrane potential-sensitive dye assay in the absence or presence of the irreversible antagonist  $\beta$ -chlornaltrexamine ( $\beta$ -CNA) to reduce receptor reserve (Fig. 1F, fig. S1A).

**Commented [AV12]:** This refers to the cartoon. Label the cartoon as panel A and panels A–E as panels B–F then relabel the callouts accordingly. See my comments in the Figure Legends.

**Commented [AV13]:** Label the cartoon as panel A and relabel panels A–E as panels B–F then adjust the callouts in the text accordingly. I already made the adjustments to the callouts, but please check to see that I got them all. See my comments in the Figure Legends.

**Commented [AV14]:** Label the cartoon as panel A and relabel panels A–E as panels B–F then adjust the callouts in the text accordingly. I already made the adjustments to the callouts, but please check to see that I got them all. See my comments in the Figure Legends.

We also constructed concentration-response curves of the nine opioid ligands using BRET assays to monitor GRK2 and  $\beta$ -arrestin 2 recruitment (Fig. 2, A–C), as well as MOR trafficking to Rab5a-positive endocytic compartments (Fig. 2D). Of note, to improve assay signal, we overexpressed GRK2 in the  $\beta$ -arrestin 2 recruitment and MOR trafficking assays (33), because we could not obtain a quantifiable response to all ligands with only endogenous amounts of GRK in these assays (fig. S1B, S1C). Altogether, we constructed concentration-response curves to accurately measure the responses the nine ligands of interest. These curves were then used to estimate signalling efficacy and biased factors.

### ***Oliceridine, PZM21, and SR-17018 have low intrinsic efficacy for receptor activation and G protein signalling***

Signaling efficacy was estimated using the operational model of agonism and quantified as the operational efficacy,  $\tau$ , a parameter comprised of both receptor density and signaling efficiency of the agonist-occupied receptor (Table 3, see Materials and Methods for detailed description of the analyses). In all G protein assays oliceridine, PZM21, and SR-17018 had lower intrinsic efficacy than did morphine, itself a partial agonist relative to fentanyl and DAMGO (Fig. 2E). Oxycodone had similar efficacy to morphine, and buprenorphine had very low efficacy. Together, these assays capture different aspects of receptor–G protein coupling and the amounts of signal amplification. In pathways with very limited signal amplification due to direct proximity to receptor activation (Nb33 recruitment, mG<sub>s1</sub> recruitment) or with very limited receptor reserve from partial irreversible antagonism (GIRK activation) partial agonists had lower efficacy and potency than in highly amplified pathways (G<sub>o12</sub> activation and AC inhibition) (Tables 1 and 2). In pathways proximal to receptor activation (Nb33 recruitment, mG<sub>s1</sub> recruitment) or with very limited receptor reserve (GIRK activation) partial agonists had lower efficacy and potency than in highly amplified, strongly coupled pathways (G<sub>o12</sub> activation and AC inhibition) (Tables 1 and 2). However, despite these differences, estimates of relative efficacy for activating G proteins were remarkably consistent across assays of receptor and G protein activation. In addition, assays of receptor regulation, including interactions between MOR and GRK2,  $\beta$ -arrestin 2, and Rab5, showed a similar pattern of agonist efficacy with rank order of maximal effect largely conserved across all assays. Oliceridine, PZM21, SR-17018, and buprenorphine displayed lower efficacy in receptor regulation than did oxycodone and morphine (Fig. 2E and Table 1).

We also monitored the kinetics of receptor-effector coupling in real-time assays of Nb33 and mG<sub>s1</sub> recruitment. SR-17018 recruited Nb33 and mG<sub>s1</sub> to the MOR at a slower rate than did morphine, DAMGO or the other putatively biased compounds oliceridine and PZM21 (fig. S2, A and B). The maximum response to SR-17018 did not occur until 5 minutes after agonist addition, whereas it occurred within 1 minute of stimulation with oliceridine, PZM21, or buprenorphine. Recruitment of Nb33 and mG<sub>s1</sub> to the MOR by oliceridine, PZM21, and SR-17018 was rapidly reversed following addition of 10  $\mu$ M the MOR antagonist naloxone, but that induced by buprenorphine was not. This was due to either the inability of naloxone to compete with the very high affinity of buprenorphine or to the extremely slow dissociation rate of this compound, or some combination of the two.

**Commented [AV15]:** As for Fig. 1, label the cartoon in Fig. 2 as panel A and relabel the panels A–E as panels B–F then adjust the callouts in the text accordingly. I've already made the adjustments to the in-text callouts, but please check to see that I got them all. See my comments in the Figure Legends.

**Commented [AV16]:** The wording of this sentence makes it unclear. The proximal pathways and pathways with limited reserve didn't have lower efficacy – pathways can't have efficacy. You must have meant that the agonists showed low efficacy in activating proximal pathways and pathways with limited reserve but high efficacy in activating pathways with signal amplification. Please reword.

You also need to explain that "pathways with limited reserve" means pathways that requires that a higher proportion of the receptors be activated in order to stimulate maximal downstream signalling, meaning that there are fewer receptors that are not involved in stimulating the response. It might be best to explain high vs. low reserve pathways in the introduction along with the concept of receptor reserve.

**Commented [AV17]:** The wording of this sentence makes it unclear. The proximal pathways and pathways with limited reserve didn't have lower efficacy – pathways can't have efficacy. You must have meant that the agonists showed low efficacy in activating proximal pathways and pathways with limited reserve but high efficacy in activating pathways with signal amplification. Please reword.

You also need to explain that "pathways with limited reserve" means pathways that requires that a higher proportion of the receptors be activated in order to stimulate maximal downstream signalling, meaning that there are fewer receptors that are not involved in stimulating the response. It might be best to explain high vs. low reserve pathways in the introduction along with the concept of receptor reserve.

**Commented [M18R17]:** I have been thinking about rewording this rather than putting more in the introduction. Please bear with me.

Agonist-induced C-terminal phosphorylation of MOR, quantified by phosphosite-specific antibodies, was in agreement with the different efficacy profiles. DAMGO, fentanyl, and methadone induced multi-site phosphorylation of MOR, as expected for high-efficacy opioids (34), whereas morphine, oxycodone, oliceridine, and PZM21 only triggered phosphorylation of Ser<sup>375</sup> (Fig. 3A) (34, 35). We have previously shown buprenorphine to have a phosphorylation profile restricted to Ser<sup>375</sup>, typical of low-efficacy agonists (34). The phosphorylation profile of SR-17018 was unusual in that even though concentrations up to 10  $\mu$ M did not produce substantial phosphorylation of residues other than Ser<sup>375</sup> up until 30 min of stimulation (Fig. 3B), the MOR phosphorylation pattern induced by SR-17018 resembled that of higher efficacy ligands at the longest incubation time point (Fig. 3A). Moreover, similar to high-efficacy agonists, this multi-site phosphorylation depended on GRK2 or GRK3 (GRK2/3), because it was blocked by incubation with the GRK2/3 inhibitor Cmpd101 (fig. S3B). Naloxone blocked the phosphorylation of MOR induced by SR-17018, oliceridine, or PZM21 (fig. S3A). Such unusual behavior for SR-17018 may also be linked to  $\beta$ -arrestin recruitment data obtained using a high temporal resolution Förster resonance-energy transfer (FRET) approach, wherein SR-17018 responses showed delayed onset of arrestin recruitment (similar to Nb33 and mG<sub>s1</sub> from fig. S2) and very slow decay upon agonist washout relative to other agonists, although this observation was difficult to consistently reproduce and warrants further investigation (fig. S4).

#### ***Efficacy for G protein pathways closely predicts efficacy in receptor regulatory pathways***

To investigate the relationship between efficacy values of the test agonists in different pathways, we calculated correlation coefficients (Pearson's  $r$ ) between  $\log(\tau)$  values (Fig. 4, fig. S5, A and B). There was robust correlation between the efficacy of test partial agonists in all pathways linked to G protein activation. The  $\log(\tau)$  obtained for G <sub>$\alpha$ 12</sub> activation significantly correlated with the  $\log(\tau)$  estimated for receptor activation by Nb33 recruitment (Fig. 4A), mG<sub>s1</sub> recruitment (Fig. 4B), cAMP inhibition (Fig. 4C), and GIRK activation (Fig. 4D, coefficient of determination  $\geq 0.79$  in all cases). Similarly, estimated efficacy of test compounds was consistent in receptor regulatory pathways. The  $\log(\tau)$  value of the agonists for  $\beta$ -arrestin2 recruitment correlated very tightly to that calculated for receptor activation (Nb33) (Fig. 4E), GRK2 recruitment (Fig. 4F), and Rab5-positive endosome trafficking (Fig. 4G). Surprisingly, efficacy for G protein pathways closely predicted efficacy in receptor regulatory pathways for all agonists, even those proposed to be biased. The  $\log(\tau)$  for  $\beta$ -arrestin recruitment significantly correlated with the  $\log(\tau)$  of G protein coupling (mG<sub>s1</sub> recruitment) (Fig. 4H) and with the activation of both G <sub>$\alpha$ 12</sub> (Fig. 4I) and GIRK channels (Fig. 4J, coefficient of determination  $\geq 0.80$ ). A less robust correlation was observed between the efficacy for  $\beta$ -arrestin recruitment and cAMP inhibition, which is likely due to the high amplification and receptor reserve of the cAMP assay that results in poorer relative efficacy estimates (Fig. 4K). Together, these results show that oliceridine, PZM21, and SR-17018 exhibited consistently low intrinsic efficacy regardless of the downstream signaling pathway that was monitored, similar to that observed for buprenorphine and lower than that for morphine and oxycodone. Fentanyl and methadone were high-efficacy agonists, similar to DAMGO, in all assays in which they were studied. For each of the family of compounds tested and across a battery of G protein and regulatory assays, relative efficacy was highly conserved.

The close correlation of efficacy across assays suggested that test agonists had similar activity in both G protein and regulatory pathways. In order to test this with a common, standard measure, we used the operational model of agonism to calculate the bias factors of all the ligands across the different signaling pathways that were assayed (see Materials and Methods). No evidence of significant bias was observed towards G protein signaling compared to  $\beta$ -arrestin 2 recruitment even after overexpression of GRK2 to ensure detectable and quantifiable responses from all ligands. The only exception was fentanyl showing a slight preference for GIRK activation (Fig. 5, fig S6A, Table S1). Unexpectedly, SR-17018, which was initially described as a ligand with an extremely high G protein bias, showed no significant bias toward or away from any G protein activation measure, with a direction of bias toward Nb33 recruitment over  $\beta$ -arrestin 2 recruitment but low confidence in that estimate. Buprenorphine showed a significant bias away from GIRK activation toward  $\beta$ -arrestin 2 recruitment.

These bias profiles were different when measured with only endogenous GRK expression (fig. S6B). Although not statistically significant, oliceridine, PZM21, and buprenorphine showed bias factors in the direction of Nb33 and  $mG_{si}$  recruitment and  $G\alpha_{i2}$  and GIRK activation over  $\beta$ -arrestin 2 recruitment in these conditions. However, poor curve fits for these extremely partial agonists in the  $\beta$ -arrestin 2 recruitment assay without kinase overexpression reduced the confidence in estimates of transduction coefficient and bias (fig. S6A).

#### ***The low intrinsic efficacy of oliceridine, PZM21, SR-17018, and buprenorphine correlates with improved therapeutic windows***

Biased agonism has been postulated to underlie the improved therapeutic indices of several opioid ligands. We measured the abilities of fentanyl, morphine, oliceridine, PZM21, SR-17018, and buprenorphine to produce anti-nociception and respiratory depression in mice. The [anti-nociceptive/antinociceptive](#) action of these compounds has been shown to be abolished in MOR-knockout animals (16, 19, 36, 37) or to be sensitive to the MOR antagonist naloxone (11). These compounds represent the full range of opioid agonist efficacy we quantified in vitro. Compounds were tested at increasing doses in a hot-plate assay (anti-nociception) and whole-body plethysmography (respiratory depression) for 180 and 240 min, respectively. Compounds were delivered sub-cutaneously, with the exception of SR-17018, which was given by intra-peritoneal injection, as in previous studies (19), due to the viscous vehicle and larger volume required.

The tested compounds produced robust anti-nociception (Fig. 6A), with the exception of SR-17018, which was dose-limited by solubility. The kinetics of the [anti-nociceptive/antinociceptive](#) effects were consistent with previously reported pharmacodynamics. Fentanyl and oliceridine had rapid onset of effect, followed by fast decay, aligning with their clinical profiles and animal studies (11). The onset and lifetime of the effect of morphine was intermediate, and although buprenorphine had similar onset to morphine, its effects on nociception were substantially longer-lasting, in line with previous research in pre-clinical and clinical settings (38). PZM21 and SR-17018 had long-lasting [anti-nociceptive/antinociceptive](#) effects, as previously reported (16, 18, 19).



The primary effect of opioid agonists on mouse respiratory function is a decrease in respiratory frequency. All tested agonists reduced respiratory frequency within the test period, albeit to varying extents (Fig. 6B). Minute volume, a combined measure of breath frequency and depth, was additionally measured over the test period and showed a very similar pattern between agonists (fig. S7). With the exception of buprenorphine, test compounds had very similar kinetics of effect between anti-nociception and respiratory depression. Buprenorphine did not have a significant effect on respiratory frequency until 230 minutes after injection at both 1 and 10 mg/kg, whereas all active anti-nociceptive/antinociceptive doses reached peak effect within 30 minutes. A more highly efficacious active metabolite of buprenorphine, norbuprenorphine, has previously been reported to contribute to respiratory depression following buprenorphine injection in mice (39, 40) and may explain this result. As previously published by Hill *et al.* (18) and in contradiction of the initial report of Manglik *et al.* (16), PZM21 induced substantial respiratory depression at anti-nociceptive/antinociceptive doses. In contrast to the study of Schmid *et al.* (19), which measured respiratory depression in restrained animals, SR-17018 induced significant respiratory depression at the highest dose tested, 30 mg/kg, although this effect was minimal compared to other agonists. The initial characterisation of SR-17018 employed restrained pulse oximetry in order to estimate respiratory frequency and tested until 60 minutes post-injection. The significant effects of SR-17018 on respiratory frequency were observed in our assay, in which animals were moving freely rather than being restrained, from 60 minutes onwards. As such, variations in the assay format or time-course may explain the differences in the observed effects. In addition, similar to the critical difference between the reports of Hill *et al.* (18) and Manglik *et al.* (16), the respiratory frequency of the vehicle-treated group in Schmid *et al.* (19) dropped substantially over the test period with approximately a 25% change in frequency, whereas the vehicle groups in the present study were more stable, changing by 15%. Comparison of the respiratory effects of different ligands at equally anti-nociceptive/antinociceptive doses revealed differences in the agonists' relative potency for the two assays. Buprenorphine, oliceridine, PZM21, and SR-17018 produced less respiratory depression than did morphine or fentanyl at equi-antinociceptive doses (Fig. 6C).

The peak effect of each dose from the hot plate and whole-body plethysmography data were then plotted in a dose-response curve (Fig. 7A). This allowed calculation of the potency,  $\log ED_{50}$ , for each drug in each behavioral assay (Fig. 7B). Peak effect was used to allow the comparison between drugs with distinct kinetic profiles. A pre-clinical model of therapeutic window was calculated by subtraction of  $\log ED_{50}$  of antinociception from respiratory depression, with error propagated. Of note, because a curve fit was not possible for SR-17018 due to solubility-limited doses, potency for antinociception was taken as minimum dose reaching a peak effect of 50% MPE. In the case of respiratory frequency depression induced by buprenorphine and SR-17018, neither of which reached 50% of the maximum effect of fentanyl, the highest tested dose was represented as a minimum lower bound of potency. Thus, the therapeutic windows of SR-17018 and buprenorphine were estimated and are shown as lower bounds (Fig. 8, A and B). Regardless, our results show that oliceridine has a larger therapeutic window in this pre-clinical model than morphine or fentanyl, with buprenorphine having the safest profile. PZM21 also appeared to have an improved therapeutic index but was less accurately quantified, whereas the window of SR-17018 was estimated.

Finally, our systematic pharmacological characterization in HEK293 cells allowed us to investigate the relationship between in vivo therapeutic window and in vitro efficacy or bias factors. We observed an inverse relationship between efficacy and therapeutic window; high-efficacy compounds, such as fentanyl, had a very narrow safety index, whereas the lowest efficacy agonist, buprenorphine, had a broad therapeutic range (Fig. 8A). Morphine, a drug less efficacious than fentanyl but more efficacious than buprenorphine, showed an intermediate safety profile. This characterization is broadly in line with clinical experience (41) and is supported by a comparison of the signaling profiles and FDA-reported adverse event frequencies of medically used opioids, which found an association between reduced efficacy and lower rates of adverse respiratory events (42). Importantly, all the so-called biased ligands (olicecidine, PZM21, and SR-17018) fell within this correlation, with efficacy and therapeutic windows falling in the range between morphine and buprenorphine. In contrast, bias factors calculated between any G protein pathway and  $\beta$ -arrestin 2 recruitment did not predict therapeutic window (Fig. 8B, fig. S8B). As noted above, the efficacy of each compound was consistent between tested assays, such that efficacy in most pathways predicted therapeutic window (fig. S8A). Comparison of agonist activity with bias plots clearly shows that all tested compounds have similar relative activity between G protein and arrestin recruitment assays (Fig. 9, fig. S9).

## Discussion

Biased agonism at the MOR has been the focus of intense research for over a decade. Early observations in  $\beta$ -arrestin 2 knockout mice prompted the hypothesis that this scaffolding protein participated in the signaling mechanisms mediating the on-target side-effects of morphine (2, 3). This gave rise to the second hypothesis that ligands able to activate G proteins without  $\beta$ -arrestin recruitment represented a safer strategy to provide analgesia (43, 44). Oliceridine, PZM21, and SR-17018 represent the most prominent examples of opioid ligands with G protein bias, although only oliceridine has been clinically tested. However, this underlying hypothesis has recently been challenged. First, respiratory depression induced through the MOR has been shown to be at least partially mediated by receptor coupling to GIRK channels through the activation of G $\beta\gamma$ -proteins (7), with neurons in several regions of the brainstem respiratory network hyperpolarised by activation of this classical,  $\beta$ -arrestin-independent MOR signalling pathway (8). This contrasts to the absence of robust physiological evidence for a  $\beta$ -arrestin signal from the MOR affecting respiratory function (45). Second, opioid side-effect profile is not improved in a knock-in mouse expressing a phosphorylation-deficient, G protein-biased MOR (5). **Third, several laboratories have been unable to repeat the primary result of reduced morphine respiratory depression in the  $\beta$ -arrestin 2 knockout (6)–(6).** ~~In this study, the systematic pharmacological characterization of these compounds in cell lines as well as in vivo allowed their direct comparison and evaluation against a panel of existing analgesics.~~

~~We~~ Here we found that efficacy robustly correlated across five assays of G protein signaling and across multiple assays of receptor regulation. The rank order of efficacy described here in heterologous expression signaling assays is consistent with other cellular and ex vivo studies performed with these ligands, including AtT20 cells and locus coeruleus (LC) neurons (46, 47). The efficacy of all test agonists, including both existing analgesics and new compounds described as biased for  $\beta$ -arrestin 2 recruitment, was closely predicted by accurate G

**Commented [AV19]:** The Discussion is a bit too long. Please try to shorten it by about 15-20%.

**Commented [MC20R19]:** We've done our best by shortening it by 10%. We feel discussion here is very important given the recent controversy and interest arisen by the recent publication of the  $\beta$ -arrestin KO data (ref. 6).

**Commented [AV21]:** We do not allow authors to cite personal communication or unpublished results. If this data has not been published or is not in press, you will need to remove this sentence.

**Commented [M22R21]:** This should have been the same as Kliever in press. It is now reference 6

protein efficacy estimates. This suggests that MOR activation by these agonists is similar for all tested assays, and indeed operational analysis found no evidence of significant agonist bias. All three putatively biased agonists had consistently low efficacy compared to morphine. Although previous descriptions of biased MOR agonists by Manglik *et al.* (16) and Schmid *et al.* (19) did not quantify efficacy explicitly, it is possible to reanalyse the cell-based assays used in these studies, which show low receptor reserve and use DAMGO as a full reference agonist to generate efficacy estimates (Table 4S2,(23)). The efficacy of oliceridine, PZM21, and SR-17018 compared to morphine estimated in this manner aligns with our own characterization of these ligands as low-efficacy agonists, as well as with other studies (18, 25, 26). ~~Work at the  $\beta_2$ -adrenergic receptor ( $\beta_2$ AR) has suggested that different active-state receptor conformations are responsible for the lower efficacy of partial agonists [46, 47]. At the MOR, a recent report also suggests differential ligand-receptor interactions between differently biased agonists with similar intrinsic efficacy (endomorphin 2 and bitorphin)[44], although these interactions have not yet been tested experimentally. However, the impact of such conformations to intracellular signalling as well as to in-vivo GPCR-mediated responses still remains unknown.~~

**Commented [MC23]:** We have deleted the following discussion as we don't think it fits here. However this was a request of one reviewer and may need to be re-inserted.

**Commented [MC24]:** In our opinion, this could be deleted. This was requested by a reviewer, but we don't really feel it fits in the flow of the discussion.

The relative potencies obtained from our concentration-response curves are also in agreement with previous reports (Table 2)(11, 16, 19, 24, 48). This is relevant because we designed each assay to quantitatively differentiate between responses to all test ligands, including very weak partial agonists such as buprenorphine. In assays with minimal amplification (~~specifically~~ recruitment of Nb33, mGsi, GRK2, and  $\beta$ -arrestin 2), which may be assumed to detect a direct, one-to-one interaction between receptor and the recruited protein,  $E_{max}$  of an agonist directly reflects efficacy (23). Indeed, this is the case upon comparison of agonist  $E_{max}$  values across these low-amplification assays. The maximal effect of each agonist was consistent between assays of recruitment (Table 1), as would be expected ~~if it were assumed that these assays detect a direct, one-to-one interaction between receptor and recruited protein. In this circumstance~~ Thus, with a lack of amplification,  $E_{max}$  values reflects efficacy directly and comparison across assays shows no evidence of agonist bias.

~~The use of analytical methods such as those based on the Black and Leff operational model of agonism have provided a useful means to quantify agonist activity. The most accepted method for detection and quantification of biased agonism is comparison of "transduction coefficients" based on the Black and Leff operational model of agonism (49). However, these existing analytical methods are not robust for assays with substantial confounding factors. The kinetics of ligand binding and signal transduction (22), receptor and effector localization (50), as well as limited assay range in systems with high amplification or extremely partial agonists, may can all substantially alter conclusions regarding agonist activity. Using experimental approaches that allow for the measurement of responses in real time we captured the peak effect of each drug for each signal transduction pathway. This also allowed us to minimise-minimize the influence of receptor desensitisation, which occurs rapidly following stimulation, on measures of G-protein activation (35). Moreover, our experimental setup~~ Our approach also ensured that the assays used were not limited by narrow signal windows. This was achieved by using irreversible receptor inactivation for GIRK channel activation and exploiting the linear receptor-effector relationship with an excess of reporter in the case of both Nb33 and mGsi recruitment assays. In doing so, we captured the true range of MOR agonist efficacy. Similarly, we adjusted our experimental conditions to minimize the "floor" effect of ligands with minimal agonist activity. ~~This was the~~

~~case in  $\beta$ -arrestin 2 recruitment and internalization assays, wherein we by overexpressed overexpression of GRK2 to ensure a detection window that fitted all ligands (33). These considerations have a profound impact on the accuracy of bias quantification.~~

When GRK2 was overexpressed, we did not observe bias for any of the compounds previously ~~characterised~~ characterized as G protein-biased (Fig. 5). Similarly, when examined under endogenous GRK expression, oliceridine, PZM21, and buprenorphine showed no significant G protein bias ~~although there was a tendency in the direction of mGsi recruitment,  $G_{\alpha i}$  activation, and inhibition of cAMP accumulation~~ (fig. S6B). Moreover, ~~and~~ as expected from the correlations between the efficacies of the ligands across different pathways (Fig. 4), no significant bias was detected when using an alternative, efficacy-driven approach for bias determination (25) (fig. S6C). Quantification of ligands that display the pronounced partial agonism of these compounds in any low-amplification pathway is confounded by poor curve fit to low responses (51). No significant bias of oliceridine and PZM21 was previously observed with operational analysis (16). For SR-17018, which was previously described to have an extremely high bias factor toward GTP $\gamma$ S and AC inhibition over  $\beta$ -arrestin 2 recruitment, we did not detect bias toward any of the G protein pathways studied. The previous study compared activity in end-point assays between membranes of CHO cells (GTP $\gamma$ S), whole CHO cells (cAMP accumulation), whole U2OS cells ( $\beta$ -arrestin 2 recruitment), and membranes from mouse brain (GTP $\gamma$ S). As described above, the current study used real-time assays on a consistent cellular background. Given the notably unique kinetics of SR-17018 across assays, capturing peak effect rather than a pre-defined end-point may partially explain discrepancies. Additionally, Schmid *et al* (2017) constrained the parameters of operational analysis to permit a curve-fit to extremely low activity compounds in the  $\beta$ -arrestin 2 recruitment assay. As shown extensively by Stahl *et al*. (51), operational curve fitting to extremely low responses produces large variability in parameter estimates.

GRK2 overexpression, matching the heterologous overexpression of MOR and  $\beta$ -arrestin 2, improves assay signal-to-noise and therefore improved the statistical confidence. Poor quantification of curve position and plateau due to assay parameters is likely partly responsible for differences between our conclusions and those of [some](#) previous studies. As discussed theoretically by Conibear & Kelly (20), intrinsic system bias may give rise to substantial differences between assays and apparent agonist bias, particularly for low-efficacy agonists compared between highly amplified and linear signaling pathways. Low-efficacy agonists, such as buprenorphine and the putatively biased compounds, will produce strong responses in highly amplified G protein assays while having little activity in no-amplification, protein-protein interaction assays such as  $\beta$ -arrestin recruitment (52, 53). Indeed, buprenorphine itself has been previously proposed to be a G protein-biased ligand (17, 54) though here and in previous studies it has consistently been observed to be an extremely partial agonist (24). This reinforces how varying signal amplification across assays can confound interpretation, most profoundly in the case of partial agonists.

~~Given the issues highlighted with operational analysis, including criticism of its application to linear systems such as arrestin recruitment [21], graphical representation of agonist activity in 'bias plots' is an accepted accessory measure for the identification of biased agonists [49, 55].~~ Equimolar comparison of each agonist's

**Commented [AV25]:** Do not use phrases like "as shown in Fig. 1B...", "are presented in Fig. 1B...", etc.. Refer to figure panels or other data elements only in parentheses.

activity in G protein activation and  $\beta$ -arrestin 2 recruitment, both with and without GRK2 overexpression, illustrates the differences in signal amplification between these assays (Fig. 9, fig. S9) (49, 55). Qualitative comparison of the location of a test agonist over this plot compared to a reference agonist is used to identify biased ligands using this method. Importantly, given the proposal of agonist bias, regardless of kinase overexpression all tested agonists produce similar levels of  $\beta$ -arrestin 2 recruitment at a given amount of G protein activation. Observation of the substantial difference in amplification between G protein activation and  $\beta$ -arrestin 2 recruitment under endogenous GRK expression illustrates how the window for detection of G protein-biased agonism in these conditions is limited. Similar results were obtained when comparing receptor activation, mGSI recruitment, cAMP inhibition, and GIRK channel activation both with and without receptor knockdown (fig. S9).

We studied the efficacy of test ligands in assays explicitly designed to capture the true range of opioid agonist signalling. These assays Our selection of opioids recapitulated the known large range of efficacy of existing opioid analgesics, from low-efficacy buprenorphine to high-efficacy fentanyl, in a wider range of functional responses (24). Buprenorphine is an opioid with unique pharmacology, having extremely high affinity with low partial agonism at MOR, antagonism of  $\kappa$ -opioid receptors (KORs), and interactions with the nociceptin/opioid-like receptor (ORL1) (56). It has been well characterized as having a 'ceiling' effect on respiration in both rodent and human studies, whereby increasing doses well past the therapeutic range do not additionally depress breathing (57), and it has been implicated in lethal overdose much less frequently than other opioids (58). Buprenorphine's low intrinsic efficacy is routinely assumed to underly this plateau of effect on human respiration, and this partial agonism did not appear to limit the clinical analgesic effect in a review of clinical trials (59). Indeed, a review from a consensus group concluded that buprenorphine produced equivalent analgesia to full opioid agonists with reduced likelihood of severe respiratory depression (60). A confounding factor in the study of this agonist, in addition to its lack of selectivity, is the production of active metabolites such as norbuprenorphine, which has lower affinity but higher efficacy than the parent compound (61). As noted above, this metabolite has been implicated in pre-clinical studies of buprenorphine respiratory depression and, in the rare cases of mortality, has been detected in blood plasma in amounts similar to that of buprenorphine (62). The accumulation of active metabolites is a potential explanation for the delayed effect of buprenorphine on respiratory depression in the present study (Fig. 6B). In any case, given that buprenorphine's anti-nociceptive effect in mice is abolished by MOR knockout (37), its superior profile here can be understood through its partial agonism at that receptor. Other opioids with very low intrinsic efficacy, including dezocine and nalbuphine, have also been shown to have ceiling effects on respiration in human studies (63, 64). A meta-analysis of randomised controlled trials concluded that nalbuphine produced equivalent analgesia to morphine (65). By contrast, the high-efficacy analgesic fentanyl has a very narrow window of therapeutic effect versus respiratory depression (66).

In the context of buprenorphine's improved profile, both clinically and in pre-clinical models, relative to that of morphine and the high-efficacy agonist fentanyl, we have identified that the therapeutic window of putatively biased MOR agonists has a potential alternative mechanism. Although still inducing similar levels of respiratory depression to morphine at high doses, oliceridine had a slightly but significantly improved window

that is in line with previous pre-clinical and clinical work (11, 67). Similarly, PZM21, though not as accurately quantified due to lower potency and solubility concerns, appeared to have an equivalent increase in window, as did the estimated profile of SR-17018. Given the challenges to the arrestin hypothesis detailed above, the low intrinsic efficacy of these compounds is a more plausible explanation for these results.

Intrinsic efficacy, as estimated in any pathway, robustly predicted the rank order of the test compounds in an in vivo model of separation between anti-nociception and respiratory depression, arguably the most clinically important opioid-induced side effect (Fig. 8A, fig. S8A). It will be of interest in future to determine if other adverse effects, such as ~~the severity of~~ constipation, tolerance, and physical dependence exhibit comparable correlations. Buprenorphine's therapeutic utility is limited by its poor bioavailability, active metabolites, and concerns around its extremely high MOR affinity, including reversibility of respiratory depression and complex pharmacodynamic and kinetic relationships. The new MOR ligands studied here may have advantages in these and other aspects that make them superior platforms for further drug development. The contribution of off-target effects, most critically interactions with KOR,  $\delta$ -opioid receptor (DOR), and ORL1, to this pre-clinical model was not directly examined here. The test agonists all have higher potency for MOR signalling over other targets, and MOR knockout abolishes the ~~anti-nociceptive/antinociceptive~~ effects of fentanyl, morphine, buprenorphine, PZM21, and SR-17018 (16, 19, 37), whereas oliceridine's effects are naloxone-sensitive (11). ~~Administration of selective DOR and KOR agonists (DPDPE and U69,593 respectively), had no significant effect on respiration in a previous study on a panel of opioids in rats [68].~~ Our data from the pre-clinical therapeutic window model therefore most likely reflects primarily MOR-mediated effects of these agonists, although off-target effects could contribute to opioid effects, particularly at high doses. Additionally, distinct pharmacokinetics have been proposed to contribute to differences in opioid respiratory depression (66) ~~although -Herehere~~ we showed that, for most agonists, the kinetics of anti-nociception and effects on respiratory frequency were consistent, ~~and that kinetically "fast" agonists did not have a consistently poorer or superior profile.~~

The phosphorylation pattern, or "barcode", of the carboxy tail of the MOR is also indicative of the ability of opioid ligands to promote  $\beta$ -arrestin recruitment and receptor internalization. It is widely accepted that high-efficacy ligands trigger robust hierarchical and sequential multisite receptor phosphorylation, whereas low-efficacy ligands only induce phosphorylation of Ser<sup>375</sup> ~~in mice~~ (Ser<sup>377</sup> in humans). Such differential phosphorylation has previously been described for DAMGO, fentanyl, oxycodone, morphine, and buprenorphine (34, 35). Our data show that the phosphorylation induced by oliceridine and PZM21, even at 10  $\mu$ M for 30 min, is limited to Ser<sup>375</sup>, supporting their classification as low-efficacy ligands. Of note, although the phosphorylation of Ser<sup>375</sup> by oliceridine had previously been reported (11), ~~its potential phosphorylation barcode had not been previously assayed.~~ The phosphorylation pattern induced by SR-17018 during the first 20 min of incubation agrees with that of a low-efficacy agonist. ~~After additional incubation, up to~~ However, after 30 min, the MOR phosphorylation induced by this ligand resembled that of the high-efficacy ligands, both in terms of the pattern as well as in terms of its sensitivity to GRK2 inhibitors. Thus, although the pharmacological ~~eharacterisation~~ ~~characterization~~ of SR-17018 in vitro and in vivo suggest that its actions could be explained by its low intrinsic activity, these latter observations suggest that an unappreciated mechanism of action of SR-17018 may still contribute to in vivo effects. Unfortunately, due to solubility concerns at high doses, the

Commented [AV26]: Correct as edited?

Commented [MC27R26]: yes

Commented [AV28]: I revised to avoid claims about being first.

Commented [MC29R28]: ok

therapeutic window of this compound had to be estimated, thus further studies on the signalling mechanisms and in vivo effects of SR-17018 are required for the full understanding of this compound's behaviour.

Taken together our results draw attention to alternative mechanisms that influence the in vivo safety profiles of opioid ligands with regards to analgesia and respiratory depression. Whereas biased agonism is one hypothesis that may rationalize opioid actions, pharmacological parameters such as low G protein efficacy may also plausibly underlie the favourable therapeutic window of new opioids and should be taken into account. Given that the existing extremely low-efficacy agonist, buprenorphine, has been shown to produce reasonable analgesia with reduced side effects and overdose liability, but is limited by a complex pharmacological profile, there is clear scope for future drug development to optimize efficacy at this target.

## Materials and Methods

### Drugs

Oliceridine (TRV130) hydrochloride was supplied by Adooq Bioscience, CA, USA as distributed by Sapphire Bioscience, Australia. Morphine hydrochloride was supplied by GlaxoSmithKline, Australia or by Hameln Inc., Hameln, Germany. Buprenorphine was supplied by the National Measurement Institute, Department of Industry, Innovation and Science, Australia. PZM21 was supplied by (MedKoo Bioscience). SR-17018 was supplied synthesised as previously described (19). Oxycodone hydrochloride was supplied by Mundipharma; levomethadone hydrochloride was supplied by Sanofi-Aventis, and fentanyl citrate was supplied by Rotexmedica. Cmpd101 was from HelloBio. All other reagents were from SigmaAldrich.

For in vivo compatible vehicle solubility, buprenorphine, PZM21 and SR-17018 were converted to hydrochloride salt form (see below). MS and NMR analysis confirmed the structures of buprenorphine, PZM21 and SR-17018 used in animal assays was as previously published.

### Plasmids

Nb33 cDNA was obtained from Mark von Zastrow and Ashish Manglik (28) and was subcloned into a Clontech pVenus vector resulting in addition of the Venus tag at the C-terminus. NES-Venus-mGsi (29), Rab5-Venus (68), Venus155-239-Gβ1, Venus1-155-Gγ2, masGRKct-RLuc8 (30, 69) were obtained from Nevin Lambert (University of Georgia, USA). FLAG-MOR-NLuc was generated by cloning FLAG-MOR (35) into a NLuc-vector from Promega. GRK2-Venus was from D. Jensen (Columbia University, New York), β-arrestin2-YFP were provided by M. Caron (Duke University, North Carolina), GRK2 WT was from M. Smit (Vrije Universiteit, Amsterdam). CAMYEL sensor was from ATCC.

### Cell culture

Human embryonic kidney 293 (HEK293) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% v/v FBS at 37°C in a humidified atmosphere of 95% air and 5% CO<sub>2</sub>. Transient transfections for BRET experiments were performed using linear polyethylenimine (PEI) with a molecular weight of 25kDa (Polysciences) at a DNA:PEI ratio of 1:6 as previously described (35). Cells were transfected in a 10 cm dish with 1 µg of donor (FLAG-mMOP-NLuc) and 4 µg of acceptor (Nb33-Venus, mGsi-Venus, GRK2-Venus, β-arrestin2-YFP or Rab5a-Venus). For G protein activation assays, cells were transfected with 2 µg of Gα<sub>i2</sub>, 1 µg of Venus155-239-Gβ1, 1 µg of Venus1-155-Gγ2 and 1 µg of masGRK3ct-RLuc8. For cAMP inhibition assays, cells were transfected with the CAMYEL BRET biosensor (2.5 µg of FLAG-mMOP and 2.5 µg of CAMYEL biosensor per dish). For GRK2 overexpression experiments, cells were co-transfected with an additional 2 µg of GRK2 WT.

The coding DNA sequence for human GIRK4 (NM\_000890.3) with a single c-myc tag inserted between GIRK4 amino acids 121V and 122G was synthesised by GenScript in pUC57. The GIRK4 construct was then subcloned into pcDNA3.1+ (ThermoFisher, MA, USA). HEK293 cells stably expressing GIRK4 channels (HEK293-GIRK4) were generated by transfecting the GIRK4-pcDNA3.1+ plasmid into HEK-Flp-in-293 cells (ThermoFisher, R75007). FLAG-mMOR in pcDNA5/FRT was subsequently stably transfected into the vacant



Flp-In site according to the manufacturer's instructions (ThermoFisher). HEK293-GIRK4-mMOR cells were grown in DMEM (Gibco ThermoFisher) supplemented with 10% v/v fetal bovine serum and penicillin/streptomycin.

#### **Bioluminescence Resonance Energy Transfer (BRET)**

Cells were replated into poly-D-lysine-coated 96-well plates 24 h after transfection and allowed to adhere overnight. BRET experiments were performed 48 h post-transfection. Furimazine (Promega) was used for Nanoluciferase (NLuc)-tagged constructs and Coelenterazine h (NanoLight Technologies, AZ) was used for Renilla luciferase (RLuc8)-tagged constructs at a final concentration of 5  $\mu$ M. Luciferase substrate was added immediately before dual fluorescence/luminescence measurement in a PHERAstar Omega plate reader (BMG LabTech). The BRET signal was calculated as the ratio of light emitted at 530 nm by YFP or Venus (optic module 535+/-30) over the light emitted at 430 nm by NLuc or RLuc8 (optic module 475+/-30). Concentration-response curves were constructed from the 10 min time point after agonist addition for all pathways except Rab5-Venus for which concentration-response curves were constructed 30 min after stimulation. To measure agonist-induced adenylate cyclase (AC) inhibition, 10  $\mu$ M forskolin was co-added with agonists to induce cAMP production. BRET signal from vehicle-treated cells was subtracted.

#### **Western blotting**

HEK293 cells stably expressing HA-MOR were seeded onto poly-L-lysine-coated 60 mm dishes and grown to 90 % confluence. After cells were either exposed or not exposed to different agonist, cells were lysed in RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1 % Nonidet P-40, 0,5 % sodium deoxycholate, 0,1 % SDS) containing protease and phosphatase inhibitors (Complete mini and PhosSTOP; Roche Diagnostics). Pierce™ HA epitope tag antibody agarose beads (Thermo Scientific) were used to enrich glycosylated proteins. To elute proteins from the beads the samples were incubated in SDS sample buffer for 25 min at 43°C. Supernatants were separated from the beads, loaded on 8% SDS polyacrylamide gels and immunoblotted onto nitrocellulose afterwards. After blocking, membranes were incubated with either anti-pT370, anti-pS375, anti-pT376 or anti-pT379 antibody, followed by detection using a chemiluminescence detection system. Blots were subsequently stripped and incubated again with the phosphorylation-independent antibody anti-HA to confirm equal loading of the gels. Protein bands on western blots were exposed to X-ray film. Films exposed in the linear range were then densitized using ImageJ 1.37v.

#### **Data analysis**

GraphPad Prism software (v. 8.0) was used for data and statistical analyses which are specifically described in the figure legends. Concentration-response curves were fitted assuming a Hill slope of unity using the following three parameter equation

$$\text{Response} = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log EC_{50} - \log[A])}}$$

Where Top and Bottom represent the maximal and minimal asymptote of the concentration-response curve, [A] is the molar concentration of agonist and EC50 is the molar concentration of agonist required to give a response which is halfway between the bottom and top responses. Data are presented as percentage of DAMGO after

normalisation to the bottom and top values of the DAMGO response. Fitting concentration-response curves with a variable Hill slope shared between agonists produced slopes not significantly different to one for most assays, and nested extra sum-of-squares F-tests preferred the simpler fixed-slope model.

For the determination of  $\tau$  values and transduction coefficients ( $\log(\tau/K_A)$ ), concentration-response data were fit to the operational model (70) - partial agonism equation:

$$Y = \text{Basal} + \frac{E_m - \text{Basal}}{1 + \left( \frac{10^{\log K_A} + 10^{\log[A]}}{10^{\log \tau} 10^{\log[A]}} \right)^n}$$

where  $E_m$  is the maximal possible response of the system, Basal is the basal level of response,  $K_A$  represents the equilibrium dissociation constant of the agonist (A) and  $\tau$  is an index of the signaling efficacy of the agonist, and  $n$  is the slope of the transducer function. The analysis assumes that the transduction machinery utilized for a given cellular pathway are the same for all agonists, such that the basal,  $E_m$  and  $n$  are shared between agonists. Within each individual experiment, basal activity was constrained to 0 and  $E_m$  was constrained to 100. The value of  $n$  was constrained to 1 as protein-protein interaction assays of Nb33, mGsi, GRK and  $\beta$ -arrestin 2 recruitment, as well as Rab5 trafficking, are linear under operational model assumptions. As noted above, test fitting of concentration-response curves to a variable Hill slope, that is  $n \neq 1$ , did not provide evidence for operational analysis with a variable value of 'n' in the remaining assays. Within each individual experiment,  $\tau$  and  $\log(\tau/K_A)$  values were calculated using DAMGO as the reference agonist. In the instances where low efficacy agonists become full agonists (i.e. in G protein activation and AC inhibition assays),  $\log K_A$  was assumed to be equal to the  $\log K_A$  value of each agonist obtained in the mGsi recruitment assay. Functional affinity of each agonist for the receptor state producing a mGsi interaction likely reflects affinity for the active, G protein coupling receptor state. In the case of the GIRK membrane-potential assay,  $\log(\tau/K_A)$  and  $\tau$  estimates were generated by simultaneous curve fitting of the untreated and  $\beta$ -CNA treated concentration-response curves.

Biased agonism was quantified as previously described (22, 71). To exclude the impact of cell-dependent and assay-dependent effects on the observed agonism at each pathway, the  $\log(\tau/K_A)$  value of a reference agonist, in this case DAMGO, was subtracted from the  $\log(\tau/K_A)$  value of the agonists of interest to give a  $\Delta\log(\tau/K_A)$  value with error propagated according to standard rules. The bias was then calculated for each agonist at two different signalling pathways by subtracting the  $\Delta\log(\tau/K_A)$  of one pathway from the second pathway to give a  $\Delta\Delta\log(\tau/K_A)$  value, with error again propagated. Degrees of freedom of the resulting variable, for calculation of 95% confidence intervals as shown in figures, and for statistical testing, was estimated by the Welch-Satterthwaite equation due to unequal variance of the fitted parameters between assays. A lack of biased agonism will result in values of  $\Delta\Delta\log(\tau/K_A)$  not significantly different from 0. When fold changes in bias are described, this was calculated by converting values of  $\Delta\Delta\log(\tau/K_A)$  to the corresponding antilog value.

Biased agonism was statistically tested by a one sample t-test comparing a given  $\Delta\Delta\log(\tau/K_A)$  value to 0, the subtracted value of the reference agonist DAMGO. P-values of this test were multiplicity corrected using the Holm-Sidak method within each set of comparisons. One sample statistical tests to 0 performed for Fig. 5 (all

agonists compared between  $\beta$ -arrestin 2 recruitment with GRK2 overexpression and Nb33/mGsi/GPA/cAMP/GIRK assays for a total of 38 comparisons) and performed for fig. S6B (all agonists compared between  $\beta$ -arrestin 2 recruitment with endogenous GRK and Nb33/mGsi/GPA/cAMP/GIRK assays for an additional, separate 38 comparisons) were multiplicity corrected within those sets of analyses.

Correlations were computed using Pearson's correlation analysis to assess whether the  $\tau$  values of each agonist in a particular pathway vary together with the  $\tau$  values of the same agonist in a different pathway. A strong correlation was reported when  $P < 0.05$  and the fraction of the variance that is "shared" between the two variables is reported as  $R^2$ .

#### **Membrane potential assay**

MOR mediated activation of potassium channels was assayed using a membrane-potential dye from Molecular Devices (FLIPR Membrane Potential Assay kit) as previously described (72). On the day before the experiment, HEK293-GIRK4-MOR cells were plated into clear bottomed, black walled 96 well plates (Corning) pre-coated with poly-D-lysine for cell adherence. Cells were plated in L-15 media (Gibco ThermoFisher) supplemented with 1% v/v FBS. The following day, cell medium was aspirated and plates were treated with either 200 nM  $\beta$ -chloralntrexamine hydrochloride (SigmaAldrich) or HEPES-buffered salt solution (HBSS) vehicle control for 20 minutes. Testing of each compound in untreated and  $\beta$ -CNA treated conditions was carried out in parallel. All experiments were carried out in duplicate. Cells were washed twice with HBSS following vehicle or  $\beta$ -CNA treatment and incubated for 1 hour in 50% L-15 media, 50% FLIPR dye reconstituted in HBSS in a final volume of 180  $\mu$ L per well. All test compounds were serially diluted in a vehicle of 1% v/v DMSO, 0.1 mg/mL bovine serum albumin (BSA in HBSS at 10x concentrations in a 96 well compound plate (Corning) such that final assay concentrations of solvent were 0.1% DMSO and 0.01 mg/mL BSA. Plates were read in a FlexStation 3 (Molecular Devices). Two-minute baselines were taken for each well before compound addition. Then 20  $\mu$ L of 10x concentrated compound were added; and read for three minutes. Peak effect was calculated by taking the minimum raw fluorescence value of the post-addition measurement period, expressed as a percentage of pre-addition baseline and vehicle corrected, and normalised to the maximal response to DAMGO without  $\beta$ -CNA treatment.

#### **Animals**

All experiments involving animals were approved by the University of Sydney Animal Ethics Committee (AEC. Protocol number K00/12-2011/3/5650). Experiments were performed under the guidelines of the Australian code of practice for the care and use of animals for scientific purposes (National Health and Medical Research Council, Australia, 7th Edition). Great care was taken to minimize animal suffering during these experiments and to reduce the number of animals used. Male C57Bl/6J mice were ordered from the Animal Resource Centre (Perth, Western Australia) and were between 6 and 9 weeks of age on the day of experimentation. Animals were housed no more than 6 to an individually ventilated cage in a temperature and humidity-controlled room with 12-hour reverse day-night cycle lighting (lights on at 20:00). Mice had free access to food and water and were allowed to acclimatise to the facility for at least 10 days before the experimental day. All animal experiments

were performed on mice between 9:00 and 19:00 hours in their dark, active period under a red light. Handling and acclimatisation of animals to the testing room occurred on at least four days before experimentation.

Doses in milligram per kilogram were calculated for the active component of the drug (free base). Morphine hydrochloride, fentanyl citrate, PZM21 hydrochloride and oliceridine hydrochloride were dissolved in saline for injection. Buprenorphine hydrochloride was dissolved in 20% polyethylene glycol 400 in saline. SR-17018 hydrochloride was dissolved in 1% v/v DMSO, 10% v/v Tween 80 and hydrochloric acid, before being diluted to a final concentration in saline and pH adjusted back to 7 with sodium hydroxide. Saline, 20% PEG and 1% DMSO/10% Tween 80 vehicles did not differentially affect either the hot-plate or whole body plethysmography assays. All injections were performed sub-cutaneously (s.c.) at a volume of 200  $\mu$ L, with the exception of SR-17018 and the 1% DMSO/10% Tween 80 vehicle which were administered intra-peritoneal (i.p.) at a volume of 400  $\mu$ L.

#### **Hot-plate assay**

Anti-nociception was tested using a 54 °C hot-plate with latency to response timed via foot-switch. Mice were placed onto the hot-plate within a plexiglass cylinder and observed at baseline and post-injection time-points of 5, 15, 30, 60, 90, 120 and 180 minutes. Hot-plate response was defined by hindpaw lift, flick, flutter or attending, or, rarely, a jump. Mice were rapidly removed from hot-plate following response or at cut-off. A maximum cut-off of 20 s was set to limit potential tissue damage. Experimenters were blind to drug or vehicle treatment group via coded syringe labelling. Responses have been expressed as percentage of maximum possible effect, by subtracting each animal's baseline latency and normalising to the cut-off latency.

#### **Whole-body plethysmography**

Respiratory side effects were tested using a commercial whole-body plethysmography (WBP) system (Buxco FinePointe system, DSI Instruments). Animals were acclimatised to one of two parallel WBP chambers on the day before the experiment for 10 - 15 minutes, and again on the morning of the experiment. While in WBP chambers, animals were shielded from view from other animals and the experimenter. Chamber air was refreshed from room air by the integrated system at a constant rate. All analysis of respiratory waveforms was completed by integrated FinePointe software. With the exception of fentanyl citrate, the respiratory effects of test compounds were examined in a rotating protocol in which 10 minute measurements were taken at a pre-injection baseline followed by 25, 65, 105 and 230 minutes post-injection. Animals settled in the WBP chambers over the first 5-minute period, and the second 5-minute period was taken for analysis. Four animals were rotated through a single chamber in this manner over the testing period. Due to the extremely rapid kinetics of fentanyl citrate, an alternative rapid rotating protocol was used. Mice were returned to the chamber immediately following injection for an initial 15 minute measurement, after which 10 minute measurements were taken at 35, 65, 95, 125 and 215 minutes post-injection, with two animals rotated through each chamber. For analysis, each animal's respiratory frequency was normalised to the pre-drug baseline as 100%.

#### **In vivo dose-response analysis**

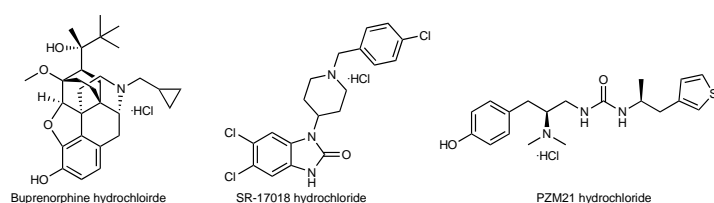
The peak response of each animal was taken for each compound for dose-response curve fitting. For comparison between hot-plate and WBP assays, only time-points captured in both assays were included in analysis. Plots were fitted to a three-parameter logistic function, with basal constrained to 0. In fitting the hot-plate assay, curve maximum was constrained to be less than the maximum possible effect. In WBP, ED<sub>50</sub> values were calculated at 50% of the maximum response to fentanyl. Therapeutic window was calculated between the two assays by subtraction of log(ED<sub>50</sub>) values to produce  $\Delta\log(\text{ED}_{50})$ . Error was propagated by standard rules, and degrees of freedom for statistical testing summed under the assumption of consistent variance. In the case of compounds that did not reach 50% of defined maximum response in the WBP assay, SR-17018 and buprenorphine, the therapeutic window was calculated by assuming that the ED<sub>50</sub> must be greater than the highest dose tested. The 95% C.I. of the  $\Delta\log(\text{ED}_{50})$  calculated for these compounds is therefore asymmetrical, with uncertainty in one direction unbounded. The maximum dose of PZM21, SR-17018 and buprenorphine in WBP were compared to the corresponding vehicle at each time point using multiple t-tests, with multiple comparisons adjusted for using the Holm-Sidak method.

#### Conversion of free base buprenorphine, SR17018, PZM21 to their monohydrochloride salt

All solvents and reagents were purchased and used from commercial sources. Nuclear magnetic resonance spectra were recorded on Bruker Advance DRX 300 at 300 MHz <sup>1</sup>H NMR frequency, chemical shifts are expressed as parts per million (ppm). All resonances are reported as chemical shift ( $\delta$ ) and are referenced to the solvent residual peak. Multiplicities are reported as follows: s (singlet), br (broad), d (doublet), dd (doublet of doublets), t (triplet), q (quartet) and m (multiplet). Coupling constants (J) are reported in Hz.

Low- and high-resolution mass spectra were obtained through electron ionisation (ESI). Low- resolution mass spectra were performed on a Finnigan LCQ mass spectrometer. High-resolution mass spectra were performed on a Bruker 7T Apex Qe Fourier Transform Ion Cyclotron resonance mass spectrometer equipped with an Apollo II ESI/APCI/MALDI Dual source.

High performance liquid chromatography (HPLC) was performed on the Waters Alliance 2695 apparatus equipped with Waters 2996 photodiode array detector, set at 254 nm. Separation using a SunFire™ C18 column (5  $\mu\text{m}$ , 2.1 x 150 mm) was achieved using water (solvent A) and acetonitrile (solvent B) at a flow rate of 0.2 mL/min. The method consisted of 0% B to 100% B over 30 minutes. HPLC data is recorded as percentage purity and retention time (RT) in minutes. All compounds showed >95% analytical purity.



A solution of free base drug (1.00 mmol) in diethyl ether (10 mL, 0.1 mM concentration) was treated with a 4 M solution of HCl in 1,4-dioxane (1.0 mL, 4.0 mmol) and stirred for 1 hour with a white precipitate formed. The solvent was removed under reduced pressure and the solid obtained dried to constant mass under high vacuum.

#### **Buprenorphine hydrochloride**

White solid. <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 6.65 (d, *J* = 8.0 Hz, 1H), 6.40 (d, *J* = 8.0 Hz, 1H), 5.85 (s, 1H), 5.36 (br. s, 1H), 4.42 (d, *J* = 0.9 Hz, 1H), 3.51 (s, 3H), 3.05 – 2.89 (m, 3H), 2.59 (dd, *J* = 10.9, 4.8 Hz, 1H), 2.42 – 2.20 (m, 4H), 2.17 (m, 1H), 1.96 (m, 1H), 1.85 – 1.67 (m, 3H), 1.35 (s, 3H), 1.32 (m, 1H), 1.06 (m, 1H), 1.04 (s, 9H), 0.81 – 0.69 (m, 2H), 0.54 – 0.43 (m, 2H), 0.10 (m, 2H), NH signal not observed; LRMS (+ESI) *m/z*: 468 (100, [M + H]<sup>+</sup>); HRMS (+ESI) *m/z*: [M + H]<sup>+</sup> calcd for C<sub>29</sub>H<sub>42</sub>NO<sub>4</sub>: 468.3108; found 468.3106.

#### **SR-17018 hydrochloride**

White solid. <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 11.25 (s, 1H), 7.55 (s, 1H), 7.40 – 7.20 (m, 4H), 7.15 (s, 1H), 6.21 (p, *J* = 3.9 Hz, 1H), 4.27 (br. s, 1 H), 3.72 (s, 2H), 3.01 (m, 2H), 2.55 (m, 2H), 2.26 (m, 2H) 1.69 (m, 2H); LRMS (+ESI) *m/z*: 410/412 (100/98, [M + H]<sup>+</sup>); HRMS (+ESI) *m/z*: [M + H]<sup>+</sup> calcd for C<sub>19</sub>H<sub>19</sub><sup>35</sup>Cl<sub>3</sub>N<sub>3</sub>O: 410.0588; found 410.0590.

#### **PZM21 hydrochloride**

White solid. <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 9.13 (s, 1H), 7.40 (dd, *J* = 4.9, 2.9 Hz, 1H), 7.09 (dd, *J* = 3.0, 1.2 Hz, 1H), 7.02 – 6.98 (m, 3H), 6.70 – 6.60 (m, 2H), 6.00 (d, *J* = 7.9 Hz, 1H), 5.55 (dd, *J* = 7.5, 2.6 Hz, 1H), 4.55 (br. s, 1 H), 3.79 (m, 1H), 3.04 (m, 1H), 2.85 – 2.60 (m, 3H), 2.62 – 2.50 (m, 2H), 2.34 (s, 6H), 2.17 (dd, *J* = 13.4, 9.3 Hz, 1H), 0.94 (d, *J* = 6.6 Hz, 3H).; LRMS (+ESI) *m/z*: 362 (100, [M + H]<sup>+</sup>); HRMS (+ESI) *m/z*: [M + H]<sup>+</sup> calcd for C<sub>19</sub>H<sub>28</sub>N<sub>3</sub>O<sub>2</sub>S: 362.1897; found 362.1897.

#### **Förster resonance energy transfer (FRET) measurements**

To measure the interaction between MOR and β-arrestin 2, HEK293T cells were transfected with 0.8 μg of MOR-YFP, 0.4 μg of human GRK2, and 0.8 μg of β-arrestin 2–mTurquoise. On the next day, cells were seeded on round 25-mm poly-d-lysine-coated coverslips, and 48 hours after transfection, FRET was measured as previously described (36)–(73), except that a light-emitting diode (LED) excitation system (pE-2; CoolLED) was used for all experiments. FRET traces were not corrected for bleaching effects.

## Supplementary Materials

**Fig. S1.** Additional concentration-response curves. Concentration-response curve for opioid ligands in GIRK activation in the absence of  $\beta$ -chlornaltrexamine treatment, and for  $\beta$ -arrestin 2 recruitment and Rab5 trafficking in the absence of GRK2 overexpression.

**Fig. S2.** Kinetics of Nb33 and mGsi recruitment and naloxone antagonism. Kinetics of Nb33 and mGsi recruitment to MOR for DAMGO, olliceridine, PZM21, SR-17018 and buprenorphine

**Fig. S3.** Sensitivity of MOR phosphorylation to naloxone and Cmpd101.

**Fig. S4.** Kinetics of interaction of MOR with  $\beta$ -arrestin 2 by FRET. Arrestin recruitment induced by DAMGO, morphine, SR-17018, PZM21 and olliceridine measured using a FRET approach

**Fig. S5.** Correlation matrix and statistical testing of correlation between  $\log(\tau)$  values. Pearson  $r$  and SEM of  $\log(\tau)$  correlations

**Fig. S6.** Additional bias calculations.  $A\log(\tau/K_A)$  for the different agonists across the different signalling pathways, biased factors  $\Delta\Delta\log(\tau/K_A)$  in the absence of GRK2 overexpression and  $\Delta\Delta\log(\tau)$  calculation of efficacy-driven bias.

**Fig. S7.** Opioid-induced respiratory depression shown analyzed as minute volume.

**Fig. S8.** Therapeutic window compared to efficacy and G-protein or  $\beta$ -arrestin 2 bias for additional signaling assays. Relationships of the therapeutic windows of these ligands with the operational efficacies or bias factors at other signalling endpoints.

**Fig. S9.** Bias plots for all additional assays. Bias plots of  $\beta$ -arrestin 2 recruitment and Nb33, mGsi, cAMP inhibition and GIRK channel activation.

Table S1. Calculation of transduction coefficients.

Table S2. G protein efficacy estimated from previous publications.

**Commented [AV30]:** After you have revised the SM, make sure that the titles listed here exactly match the title of each item as it appears in the SM file. In general, these titles are much too long. Please shorten the titles.

## References and Notes

1. Williams, J. T., et al. Regulation of  $\mu$ -opioid receptors: desensitization, phosphorylation, internalization, and tolerance. *Pharmacol Rev* 2013;65(1):p.223-254. J. Christie, Regulation of  $\mu$ -Opioid Receptors: Desensitization, Phosphorylation, Internalization, and Tolerance. *Pharmacol Rev* 65, 223-254 (2012); published online EpubDec 20 (10.1124/pr.112.005942).
2. L. M. Bohn, R. J. Lefkowitz, R. R. Gainetdinov, K. Peppel, M. G. Caron, F. T. Lin, Enhanced morphine analgesia in mice lacking beta-arrestin 2. *Science* 286, 2495-2498 (1999); published online EpubDec 24 (10.1126/science.286.5449.2495).
3. K. M. Raehal, J. K. Walker, L. M. Bohn, Morphine side effects in beta-arrestin 2 knockout mice. *J Pharmacol Exp Ther* 314, 1195-1201 (2005); published online EpubSep (10.1124/jpet.105.087254).
4. J. D. Violin, A. L. Crombie, D. G. Soergel, M. W. Lark, Biased ligands at G-protein-coupled receptors: promise and progress. *Trends Pharmacol Sci* 35, 308-316 (2014); published online EpubJul 01 (10.1016/j.tips.2014.04.007).
5. A. Kliewer, F. Schmiedel, S. Sianati, A. Bailey, J. T. Bateman, E. S. Levitt, J. T. Williams, M. J. Christie, S. Schulz, Phosphorylation-deficient G-protein-biased  $\mu$ -opioid receptors improve analgesia and diminish tolerance but worsen opioid side effects. *Nat Commun* 10, 367 (2019); published online EpubFeb 21 (10.1038/s41467-018-08162-1).
6. A. Kliewer, A. Gillis, R. Hill, F. Schmidel, C. Bailey, E. Kelly, G. Henderson, M. J. Christie, S. Schulz, Morphine-induced respiratory depression is independent of beta-arrestin2 signalling. *Br J Pharmacol*, (2020); published online EpubFeb 12 (10.1111/bph.15004).
7. G. Montandon, J. Ren, N. C. Victoria, H. Liu, K. Wickman, J. J. Greer, R. L. Horner, G-protein-gated Inwardly Rectifying Potassium Channels Modulate Respiratory Depression by Opioids. *Anesthesiology* 124, 641-650 (2016); published online EpubApr (10.1097/ALN.0000000000000984).
8. E. S. Levitt, A. P. Abdala, J. F. Paton, J. M. Bissonnette, J. T. Williams,  $\mu$  opioid receptor activation hyperpolarizes respiratory-controlling Kolliker-Fuse neurons and suppresses post-inspiratory drive. *J Physiol* 593, 4453-4469 (2015); published online EpubOct 1 (10.1113/JP270822).
9. J. J. Galligan, H. I. Akbarali, Molecular physiology of enteric opioid receptors. *Am J Gastroenterol Suppl* 2, 17-21 (2014); published online EpubSep 10 (10.1038/ajgsup.2014.5).
10. M. Z. Imam, A. Kuo, S. Ghassabian, M. T. Smith, Progress in understanding mechanisms of opioid-induced gastrointestinal adverse effects and respiratory depression. *Neuropharmacology* 131, 238-255 (2018); published online EpubMar 15 (10.1016/j.neuropharm.2017.12.032).
11. S. M. DeWire, D. S. Yamashita, D. H. Rominger, G. Liu, C. L. Cowan, T. M. Graczyk, X. T. Chen, P. M. Pitis, D. Gotchev, C. Yuan, M. Koblish, M. W. Lark, J. D. Violin, A G Protein-Biased Ligand at the  $\mu$ -Opioid Receptor Is Potently Analgesic with Reduced Gastrointestinal and Respiratory Dysfunction Compared with Morphine. *J Pharmacol Exp Ther* 344, 708-717 (2013); published online EpubMar 19 (10.1124/jpet.112.201616).
12. A. A. Altarifi, B. David, K. H. Muchhala, B. E. Blough, H. Akbarali, S. S. Negus, Effects of acute and repeated treatment with the biased  $\mu$  opioid receptor agonist TRV130 (oliceclidine) on measures of antinociception, gastrointestinal function, and abuse liability in rodents. *J Psychopharmacol* 31, 730-739 (2017); published online EpubMay 18 (10.1097/j.pain.0000000000000363).
13. C. A. Zamarripa, S. R. Edwards, H. N. Qureshi, J. N. Yi, B. E. Blough, K. B. Freeman, The G-protein biased  $\mu$ -opioid agonist, TRV130, produces reinforcing and antinociceptive effects that are comparable to oxycodone in rats. *Drug and Alcohol Dependence*, 1-24 (2018); published online EpubSep 18 (10.1016/j.drugalcdep.2018.08.002).
14. N. Singla, H. Minkowitz, D. Soergel, D. Burt, R. A. Subach, M. Salamea, M. Fossler, F. Skobieranda, A randomized, Phase IIb study investigating oliceridine (TRV130), a novel  $\mu$ -receptor G-protein pathway selective ( $\mu$ -GPS) modulator, for the management of moderate to severe acute pain following abdominoplasty. *J Pain Res* Volume 10, 2413-2424 (2017)10.2147/JPR.S137952).
15. N. K. Singla, F. Skobieranda, D. G. Soergel, M. Salamea, D. A. Burt, M. A. Demitrack, E. R. Viscusi, APOLLO-2: A Randomized, Placebo and Active-Controlled Phase III Study Investigating Oliceridine (TRV130), a G Protein-Biased Ligand at the  $\mu$ -Opioid Receptor, for Management of Moderate to Severe Acute Pain Following Abdominoplasty. *Pain Pract*, (2019); published online EpubJun 4 (10.1111/papr.12801).
16. A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hübner, X.-P. Huang, M. F. Sassano, P. M. Giguère, S. Löber, D. Duan, G. Scherrer, B. K. Kobilka, P. Gmeiner, B. L. Roth, B. K. Shoichet, Structure-based discovery of opioid analgesics with reduced side effects. *Nature* 537, 185-190 (2016); published online EpubSep 08 (10.1038/nature19112).

**Commented [AV31]:** We do not use "et al." in references. Use complete author lists instead. Please reformat your references so that they conform to Science Signaling style. If your reference manager does not contain the Sci. Signal. format, please use the J. Immunol. format.

**Commented [MC32R31]:** Done, I used the Science Signaling format downloaded from Endnote site



17. A. T. Ehrlich, M. Semache, F. Gross, D. F. Da Fonte, L. Runtz, C. Colley, A. Mezni, C. Le Gouill, V. Lukashova, M. Hogue, E. Darcq, M. Bouvier, B. L. Kieffer, Biased Signaling of the Mu Opioid Receptor Revealed in Native Neurons. *iScience* **14**, 47-57 (2019); published online EpubMay 26 (10.1016/j.isci.2019.03.011).
18. R. Hill, A. Disney, A. Conibear, K. Sutcliffe, W. Dewey, S. Husbands, C. Bailey, E. Kelly, G. Henderson, The novel  $\mu$ -opioid receptor agonist PZM21 depresses respiration and induces tolerance to antinociception. *Br J Pharmacol* **175**, 2653-2661 (2018); published online EpubJun 14 (10.1124/jpet.116.238329).
19. C. L. Schmid, N. M. Kennedy, N. C. Ross, K. M. Lovell, Z. Yue, J. Morgenweck, M. D. Cameron, T. D. Bannister, L. M. Bohn, Bias Factor and Therapeutic Window Correlate to Predict Safer Opioid Analgesics. *Cell* **171**, 1165-1170.e1113 (2017); published online EpubNov 16 (10.1016/j.cell.2017.10.035).
20. A. E. Conibear, E. Kelly, A biased view of  $\mu$  opioid receptors? *Mol Pharmacol*, mol.119.115956 (2019); published online EpubJul 07 (10.1124/mol.119.115956).
21. J. W. Black, P. Leff, Operational Models of Pharmacological Agonism. *Proceedings of the Royal Society B: Biological Sciences* **220**, 141-162 (1983); published online EpubDec 22 (10.1098/rspb.1983.0093).
22. C. Klein Herenbrink, D. A. Sykes, P. Donthamsetti, M. Canals, T. Coudrat, J. Shonberg, P. J. Scammells, B. Capuano, P. M. Sexton, S. J. Charlton, J. A. Javitch, A. Christopoulos, J. R. Lane, The role of kinetic context in apparent biased agonism at GPCRs. *Nat Commun* **7**, 10842 (2016); published online EpubFeb 24 (10.1038/ncomms10842).
23. E. Kelly, Efficacy and ligand bias at the  $\mu$ -opioid receptor. *Br J Pharmacol* **169**, 1430-1446 (2013); published online EpubJul 12 (10.1124/mol.107.039842).
24. J. McPherson, G. Rivero, M. Baptist, J. Llorente, S. Al-Sabah, C. Krasel, W. L. Dewey, C. P. Bailey, E. M. Rosethorne, S. J. Charlton, G. Henderson, E. Kelly,  $\mu$ -Opioid Receptors: Correlation of Agonist Efficacy for Signalling with Ability to Activate Internalization. *Mol Pharmacol* **78**, 756-766 (2010); published online EpubSep 17 (10.1124/mol.110.066613).
25. J. Burgueno, M. Pujol, X. Monroy, D. Roche, M. J. Varela, M. Merlos, J. Giraldo, A Complementary Scale of Biased Agonism for Agonists with Differing Maximal Responses. *Sci Rep* **7**, 15389 (2017); published online EpubNov 13 (10.1038/s41598-017-15258-z).
26. Y. Yudin, T. Rohacs, The G protein-biased agents PZM21 and TRV130 are partial agonists of  $\mu$ -opioid receptor-mediated signaling to ion channels. *Br J Pharmacol*, bph.14702 (2019); published online EpubJun 09 (10.1111/bph.14702).
27. G. L. Thompson, J. R. Lane, T. Coudrat, P. M. Sexton, A. Christopoulos, M. Canals, Systematic analysis of factors influencing observations of biased agonism at the mu-opioid receptor. *Biochem Pharmacol* **113**, 70-87 (2016); published online EpubAug 01 (10.1016/j.bcp.2016.05.014).
28. M. Stoerber, D. Jullié, B. T. Lobingier, T. Laeremans, J. Steyaert, P. W. Schiller, A. Manglik, M. von Zastrow, A Genetically Encoded Biosensor Reveals Location Bias of Opioid Drug Action. *Neuron*, 1-20 (2018); published online EpubJun 04 (10.1016/j.neuron.2018.04.021).
29. Q. Wan, N. Okashah, A. Inoue, R. Nehme, B. Carpenter, C. G. Tate, N. A. Lambert, Mini G protein probes for active G protein-coupled receptors (GPCRs) in live cells. *J Biol Chem* **293**, 7466-7473 (2018); published online EpubMay 11 (10.1074/jbc.RA118.001975).
30. I. Masuho, K. A. Martemyanov, N. A. Lambert, Monitoring G protein activation in cells with BRET. *Methods Mol Biol* **1335**, 107-113 (2015)10.1007/978-1-4939-2914-6\_8.
31. C. Galès, J. J. Van Durm, S. Schaak, S. Pontier, Y. Percherancier, M. Audet, H. Paris, M. Bouvier, Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes. *Nature Structural and Molecular Biology* **13**, 778-786 (2006); published online EpubAug 13 (10.1006/smns.1998.0125).
32. L. I. Jiang, J. Collins, R. Davis, K.-M. Lin, D. DeCamp, T. Roach, R. Hsueh, R. A. Rebres, E. M. Ross, R. Taussig, I. Fraser, P. C. Sternweis, Use of a cAMP BRET Sensor to Characterize a Novel Regulation of cAMP by the Sphingosine 1-Phosphate/G 13Pathway. *J Biol Chem* **282**, 10576-10584 (2007); published online EpubApr 30 (10.1042/bj2810073).
33. J. Zhang, S. S. Ferguson, L. S. Barak, S. R. Bodduluri, S. A. Laporte, P. Y. Law, M. G. Caron, Role for G protein-coupled receptor kinase in agonist-specific regulation of mu-opioid receptor responsiveness. *Proc Natl Acad Sci U S A* **95**, 7157-7162 (1998); published online EpubJun 9 (10.1073/pnas.95.12.7157).
34. S. Just, S. Illing, M. Trester-Zedlitz, E. K. Lau, S. J. Kotowski, E. Miess, A. Mann, C. Doll, J. C. Trinidad, A. L. Burlingame, M. von Zastrow, S. Schulz, Differentiation of Opioid Drug Effects by Hierarchical Multi-Site Phosphorylation. *Mol Pharmacol* **83**, 633-639 (2013); published online EpubMar 18 (10.1124/mol.112.082875).

35. E. Miess, A. B. Gondin, A. Yousuf, R. Steinborn, N. Mösslein, Y. Yang, M. Göldner, J. G. Ruland, M. Bünemann, C. Krasel, M. J. Christie, M. L. Halls, S. Schulz, M. Canals, Multisite phosphorylation is required for sustained interaction with GRKs and arrestins during rapid  $\mu$ -opioid receptor desensitization. *Science signaling* **11**, (2018); published online EpubJul 17 (10.1126/scisignal.aas9609).
36. H. W. Matthes, R. Maldonado, F. Simonin, O. Valverde, S. Slowe, I. Kitchen, K. Befort, A. Dierich, M. Le Meur, P. Dollé, E. Tzavara, J. Hanoune, B. P. Roques, B. L. Kieffer, Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the mu-opioid-receptor gene. *Nature* **383**, 819-823 (1996); published online EpubOct 31 (10.1038/383819a0).
37. S. Ide, M. Minami, M. Satoh, G. R. Uhl, I. Sora, K. Ikeda, Buprenorphine antinociception is abolished, but naloxone-sensitive reward is retained, in mu-opioid receptor knockout mice. *Neuropsychopharmacol* **29**, 1656-1663 (2004); published online EpubSep (10.1038/sj.npp.1300463).
38. A. Elkader, B. Sproule, Buprenorphine: clinical pharmacokinetics in the treatment of opioid dependence. *Clin Pharmacokinet* **44**, 661-680 (2005)10.2165/00003088-200544070-00001).
39. S. M. Brown, M. Holtzman, T. Kim, E. D. Kharasch, Buprenorphine metabolites, buprenorphine-3-glucuronide and norbuprenorphine-3-glucuronide, are biologically active. *Anesthesiology* **115**, 1251-1260 (2011); published online EpubDec (10.1097/ALN.0b013e318238fea0).
40. H. Alhaddad, S. Cisternino, X. Decleves, N. Tournier, J. Schlatter, F. Chiadmi, P. Risede, M. Smirnova, C. Besengez, J. M. Scherrmann, F. J. Baud, B. Megarbane, Respiratory toxicity of buprenorphine results from the blockage of P-glycoprotein-mediated efflux of norbuprenorphine at the blood-brain barrier in mice. *Crit Care Med* **40**, 3215-3223 (2012); published online EpubDec (10.1097/CCM.0b013e318265680a).
41. R. F. Wolff, D. Aune, C. Truyers, A. V. Hernandez, K. Misso, R. Riemsma, J. Kleijnen, Systematic review of efficacy and safety of buprenorphine versus fentanyl or morphine in patients with chronic moderate to severe pain. *Curr Med Res Opin* **28**, 833-845 (2012); published online EpubMay (10.1185/03007995.2012.678938).
42. B. Benredjem, J. Gallion, D. Pelletier, P. Dallaire, J. Charbonneau, D. Cawkill, K. Nagi, M. Gosink, V. Lukasheva, S. Jenkinson, Y. Ren, C. Soms, B. Murat, E. Van Der Westhuizen, C. Le Gouill, O. Lichtarge, A. Schmidt, M. Bouvier, G. Pineyro, Exploring use of unsupervised clustering to associate signaling profiles of GPCR ligands to clinical response. *Nat Commun* **10**, 4075 (2019); published online EpubSep 9 (10.1038/s41467-019-11875-6).
43. E. R. Siuda, R. Carr, 3rd, D. H. Rominger, J. D. Violin, Biased mu-opioid receptor ligands: a promising new generation of pain therapeutics. *Curr Opin Pharmacol* **32**, 77-84 (2017); published online EpubFeb (10.1016/j.coph.2016.11.007).
44. A. Madariaga-Mazon, A. F. Marmolejo-Valencia, Y. Li, L. Toll, R. A. Houghten, K. Martinez-Mayorga, Mu-Opioid receptor biased ligands: A safer and painless discovery of analgesics? *Drug Discov Today* **22**, 1719-1729 (2017); published online EpubNov (10.1016/j.drudis.2017.07.002).
45. K. M. Raehal, L. M. Bohn, beta-arrestins: regulatory role and therapeutic potential in opioid and cannabinoid receptor-mediated analgesia. *Handb Exp Pharmacol* **219**, 427-443 (2014)10.1007/978-3-642-41199-1\_22).
46. Z. Dekan, S. Sianati, A. Yousuf, K. J. Sutcliffe, A. Gillis, C. Mallet, P. Singh, A. H. Jin, A. M. Wang, S. A. Mohammadi, M. Stewart, R. Ratnayake, F. Fontaine, E. Lacey, A. M. Piggott, Y. P. Du, M. Canals, R. B. Sessions, E. Kelly, R. J. Capon, P. F. Alewood, M. J. Christie, A tetrapeptide class of biased analgesics from an Australian fungus targets the micro-opioid receptor. *Proc Natl Acad Sci U S A*, (2019); published online EpubOct 14 (10.1073/pnas.1908662116).
47. M. Sadeghi, T. M. Tzschentke, M. J. Christie, mu-Opioid receptor activation and noradrenaline transport inhibition by tapentadol in rat single locus coeruleus neurons. *Br J Pharmacol* **172**, 460-468 (2015); published online EpubJan (10.1111/bph.12566).
48. K. E. Livingston, J. R. Traynor, Disruption of the Na<sup>+</sup> ion binding site as a mechanism for positive allosteric modulation of the mu-opioid receptor. *Proc Natl Acad Sci U S A* **111**, 18369-18374 (2014); published online EpubDec 23 (10.1124/pr.54.2.323).
49. T. Kenakin, A. Christopoulos, Measurements of ligand bias and functional affinity. 1-2 (2013); published online EpubJun 17 (10.1038/nrd3954-c2).
50. B. T. Lobingier, M. von Zastrow, When trafficking and signaling mix: How subcellular location shapes G protein-coupled receptor activation of heterotrimeric G proteins. *Traffic* **20**, 130-136 (2019); published online EpubFeb (10.1111/tra.12634).

51. E. L. Stahl, L. Zhou, F. J. Ehlert, L. M. Bohn, A Novel Method for Analyzing Extremely Biased Agonism at G Protein-Coupled Receptors. *Mol Pharmacol* **87**, 866-877 (2015); published online EpubApr 23 (10.1124/mol.114.096503).
52. E. Kelly, Ligand bias at the  $\mu$ -opioid receptor. *Biochem Soc Trans* **41**, 218-224 (2013); published online EpubMar 01 (10.1523/JNEUROSCI.4874-10.2011).
53. S. A. Nickolls, A. Waterfield, R. E. Williams, R. A. Kinloch, Understanding the effect of different assay formats on agonist parameters: a study using the micro-opioid receptor. *J Biomol Screen* **16**, 706-716 (2011); published online EpubAug (10.1177/1087057111406548).
54. M. F. Pedersen, T. M. Wrobel, E. Marcher-Rorsted, D. S. Pedersen, T. C. Moller, F. Gabriele, H. Pedersen, D. Matosiuk, S. R. Foster, M. Bouvier, H. Brauner-Osborne, Biased agonism of clinically approved  $\mu$ -opioid receptor agonists and TRV130 is not controlled by binding and signaling kinetics. *Neuropharmacology*, 107718 (2019); published online EpubJul 24 (10.1016/j.neuropharm.2019.107718).
55. S. Rajagopal, S. Ahn, D. H. Rominger, W. Gowen-MacDonald, C. M. Lam, S. M. DeWire, J. D. Violin, R. J. Lefkowitz, Quantifying Ligand Bias at Seven-Transmembrane Receptors. *Mol Pharmacol* **80**, 367-377 (2011); published online EpubAug 19 (10.1124/mol.111.072801).
56. H. G. Kress, Clinical update on the pharmacology, efficacy and safety of transdermal buprenorphine. *Eur J Pain* **13**, 219-230 (2009); published online EpubMar (10.1016/j.ejpain.2008.04.011).
57. A. Dahan, Opioid-induced respiratory effects: new data on buprenorphine. *Palliat Med* **20 Suppl 1**, s3-8 (2006).
58. M. Fatseas, M. Auriacombe, Why buprenorphine is so successful in treating opiate addiction in France. *Curr Psychiatry Rep* **9**, 358-364 (2007); published online EpubOct (
59. R. B. Raffa, M. Haidery, H. M. Huang, K. Kalladeen, D. E. Lockstein, H. Ono, M. J. Shope, O. A. Sowunmi, J. K. Tran, J. V. Pergolizzi, Jr., The clinical analgesic efficacy of buprenorphine. *J Clin Pharm Ther* **39**, 577-583 (2014); published online EpubDec (10.1111/jcpt.12196).
60. J. Pergolizzi, A. M. Aloisi, A. Dahan, J. Filitz, R. Langford, R. Likar, S. Mercadante, B. Morlion, R. B. Raffa, R. Sabatowski, P. Sacerdote, L. M. Torres, A. A. Weinbroum, Current knowledge of buprenorphine and its unique pharmacological profile. *Pain Pract* **10**, 428-450 (2010); published online EpubSep-Oct (10.1111/j.1533-2500.2010.00378.x).
61. P. Huang, G. B. Kehner, A. Cowan, L. Y. Liu-Chen, Comparison of pharmacological activities of buprenorphine and norbuprenorphine: norbuprenorphine is a potent opioid agonist. *J Pharmacol Exp Ther* **297**, 688-695 (2001); published online EpubMay (
62. H. K. Kim, M. Smiddy, R. S. Hoffman, L. S. Nelson, Buprenorphine may not be as safe as you think: a pediatric fatality from unintentional exposure. *Pediatrics* **130**, e1700-1703 (2012); published online EpubDec (10.1542/peds.2012-1342).
63. A. Romagnoli, A. S. Keats, Ceiling respiratory depression by dezocine. *Clin Pharmacol Ther* **35**, 367-373 (1984); published online EpubMar (10.1038/clpt.1984.45).
64. I. D. Klepper, M. Rosen, M. D. Vickers, W. W. Mapleson, Respiratory function following nalbuphine and morphine in anaesthetized man. *Br J Anaesth* **58**, 625-629 (1986); published online EpubJun (10.1093/bja/58.6.625).
65. Z. Zeng, J. Lu, C. Shu, Y. Chen, T. Guo, Q. P. Wu, S. L. Yao, P. Yin, A comparison of nalbuphine with morphine for analgesic effects and safety : meta-analysis of randomized controlled trials. *Sci Rep* **5**, 10927 (2015); published online EpubJun 3 (10.1038/srep10927).
66. A. Yassen, E. Olofsen, J. Kan, A. Dahan, M. Danhof, Pharmacokinetic-pharmacodynamic modeling of the effectiveness and safety of buprenorphine and fentanyl in rats. *Pharm Res* **25**, 183-193 (2008); published online EpubJan (10.1007/s11095-007-9440-z).
67. E. R. Viscusi, F. Skobieranda, D. G. Soergel, E. Cook, D. A. Burt, N. Singla, APOLLO-1: a randomized placebo and active-controlled phase III study investigating oliceridine (TRV130), a G protein-biased ligand at the  $\mu$ -opioid receptor, for management of moderate-to-severe acute pain following bunionectomy. *J Pain Res* **12**, 927-943 (2019)10.2147/JPR.S171013).
68. T.-H. Lan, Q. Liu, C. Li, G. Wu, N. A. Lambert, Sensitive and High Resolution Localization and Tracking of Membrane Proteins in Live Cells with BRET. *Traffic* **13**, 1450-1456 (2012); published online EpubAug 05 (10.1111/j.1600-0854.2012.01401.x).
69. I. Masuho, K. A. Martemyanov, N. A. Lambert. (Springer New York, New York, NY, 2015), vol. 1335, pp. 107-113.
70. J. W. Black, P. Leff, Operational models of pharmacological agonism. *Proceedings of the Royal Society of London. Series B, Biological sciences* **220**, 141-162 (1983); published online EpubDec 22 (

71. T. Kenakin, C. Watson, V. Muniz-Medina, A. Christopoulos, S. Novick, A simple method for quantifying functional selectivity and agonist bias. *ACS Chem Neurosci* **3**, 193-203 (2012); published online EpubMar 21 (10.1021/cn200111m).
72. A. Knapman, M. Santiago, Y. P. Du, P. R. Bennallack, M. J. Christie, M. Connor, A Continuous, Fluorescence-based Assay of  $\mu$ -Opioid Receptor Activation in AtT-20 Cells. *J Biomol Screen* **18**, 269-276 (2013); published online EpubApr (10.1186/1471-2164-12-14).
73. V. Wolters, C. Krasel, J. Brockmann, M. Bunemann, Influence of G q on the Dynamics of M3-Acetylcholine Receptor-G-Protein-Coupled Receptor Kinase 2 Interaction. *Mol Pharmacol* **87**, 9-17 (2014); published online EpubNov 19 (10.1124/mol.114.094722).

**Acknowledgments:** We thank Mark von Zastrow and Aashish Manglik (UCSF) for providing Nb33, Nevin Lambert (Georgia U.) for providing mGsi and Nicholas Kinson for the synthesis of SR-17018. **Funding:** Program Grant (1072113 to M.J.C), Senior Principal Research Fellowship (M.J.C), Project Grant (1121029 to M.C and M.L.H) from the National Health and Medical Research Council of Australia. Auckland Medical Research Foundation Grant to N. L. G. **Author contributions:** A.G and A.B.G performed the majority of experiments, ~~analysed~~ analyzed data and wrote the manuscript. A.K., S. F and F.S. performed MOR phosphorylation assays. J.S and H.D.L performed cAMP and arrestin BRET assays. C. A. performed animal experiments. P.M, M.S performed membrane potential experiments. T.K. and T.R performed salt conversions. N.L.G generated HEK293-GIRK4 cell lines. M.K supervised salt conversions. B.K supervised synthesis of SR-17018. C.K. performed FRET experiments. M.L.H. supervised BRET experiments. M.C. supervised membrane potential experiments. J.R.L provided critical advice on data analysis. S.S. supervised phosphorylation assays. M.J.C and M.C. conceived and supervised the project, performed data analysis and wrote the manuscript. All authors reviewed the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

**Commented [AV33]:** Make sure the Funding section is complete. Authors omitting one or more funding sources is a common cause for issuing errata, so it is better to double-check these now instead of having to issue an erratum later.

**Commented [AV34]:** This statement is required. Add information about any MTAs here.

## Figure Legends

### Fig. 1. Oliceridine, PZM21, and SR-17018 have low intrinsic efficacy for receptor activation and G protein signaling. (A) Schematic showing the signalling pathways induced by MOR agonists (star)

investigated. Recruitment of  $G_{\alpha i}$  leads to inhibition of AC and therefore a reduction in cAMP. (B to F) Concentration-response curves for DAMGO, fentanyl, methadone, morphine, oxycodone, oliceridine, PZM21, SR-17018, and buprenorphine were determined for recruitment of the conformationally selective nanobody Nb33 (B), recruitment of mini  $G_i$ -protein ( $mG_{\alpha i}$ ) (C), activation of  $G_{\alpha_{i2}}$  (D), inhibition of Fsk-induced cAMP accumulation (E), and GIRK channel activation upon inactivation of MOR with  $\beta$ -CNACNA (see fig. S1 for GIRK channel activation in the absence of  $\beta$ -CAN) (F). All responses were normalized to the response induced by 10  $\mu$ M DAMGO. Data are means  $\pm$  SEM of three to ten experiments, each performed in duplicate (F) or triplicate (B–E).

Commented [AV35]: Label the cartoon as panel A. Relabel panels A–E as B–F. Change the callouts in the text accordingly.

Commented [AV36]: Please label MOR.

Commented [AV37]: Since MOR stimulates GIRK through  $G\beta\gamma$  recruitment, shouldn't CAN reduce GIRK activation? I recommend you explain this mechanism more clearly in the text.

### Fig. 2. Low intrinsic-efficacy opioid ligands show limited recruitment of regulatory proteins and internalization. (A) Schematic showing the regulatory pathways investigated. Upon receptor activation, MOR is phosphorylated by GRK2, prompting the recruitment of $\beta$ -arrestin 2 to the receptor and initiating receptor internalization towards early endosomes. XXXX. (B to D) Concentration-response curves for DAMGO, fentanyl, methadone, morphine, oxycodone, oliceridine, PZM21, SR-17018, and buprenorphine were constructed for GRK2 recruitment (B), $\beta$ -arrestin 2 recruitment (C), and MOR trafficking to Rab5-positive endosomes (D). All responses were normalised to the response induced by 10 $\mu$ M DAMGO. Data are means $\pm$ SEM of three to fourteen experiments, each performed in triplicate. (E) The operational efficacy $\log(\tau)$ for the partial agonists was calculated and plotted as a heatmap.

Commented [AV38]: Label the cartoon as panel A. Relabel panels A–D as B–E. Change the callouts in the text accordingly.

Commented [AV39]: Label the MOR that is at the membrane (no need to label the one in the endosome as well).

Commented [AV40]: Add a sentence to explain the contributions of GRK2 and bArr2 to MOR internalization into RAB5+ endosomes.

**Fig. 3. Receptor phosphorylation barcode correlates with efficacy profiles.** (A) HEK293 cells stably expressing HA-tagged MOR were stimulated with increasing concentrations of fentanyl, methadone, morphine, oxycodone, oliceridine, PZM21, or SR-17018 for 30 min. Cells were lysed and immunoblotted for MOR phosphorylated at Thr<sup>370</sup> (pT370), Ser<sup>375</sup> (pS375), Thr<sup>376</sup> (pT376), and Thr<sup>379</sup> (pT379). Blots were stripped and reprobed with an antibody against HA to confirm equal loading of HA-MOR. (B) HEK293 cells stably expressing HA-MOR were stimulated with 10  $\mu$ M DAMGO or 10  $\mu$ M SR-17018 and phosphorylation of Thr<sup>370</sup>, Ser<sup>375</sup>, Thr<sup>376</sup>, and Thr<sup>379</sup> was assessed at different time points. Blots are representative of three independent experiments.

### Fig. 4. Correlations between efficacies of opioid ligands across multiple signaling pathways. (A to D)

Correlation of the operational efficacy,  $\log(\tau)$ , for  $G_{\alpha_{i2}}$  activation by the partial agonists morphine, oxycodone, oliceridine, PZM21, SR-17018, and buprenorphine, as calculated from the concentration-response curves (Figs. 1 and 2, Materials and Methods), with the  $\log(\tau)$  obtained for receptor activation, as measured by Nb33 recruitment (A) and for recruitment of  $mG_{\alpha i}$  (B), cAMP inhibition (C), and  $G_{\beta\gamma}$ -mediated GIRK activation (D). (E to G) Correlation of the operational efficacy,  $\log(\tau)$ , calculated for  $\beta$ -arrestin 2 recruitment with the  $\log(\tau)$  obtained for Nb33 recruitment (E), GRK2 recruitment (F), and MOR trafficking to Rab5-positive endosomes (G). (H to K) Correlation of the  $\log(\tau)$  for  $\beta$ -arrestin 2 recruitment with that of G protein-mediated signaling

Commented [AV41]: Remove the boxes around panels A–D and E–K. It is OK to leave the text reading "G protein activation" and " $\beta$ -arrestin 2 recruitment," but remove the hyphen from "G-protein."

pathways, including  $mG_{si}$  recruitment (H),  $G\alpha_{i2}$  (I) and GIRK channel (J) activation, and, to a lesser degree, with inhibition of Fsk-induced cAMP accumulation (K).

**Fig. 5. Bias factor calculations for opioid ligands.** Bias factors of all opioids were calculated using DAMGO as a reference (see Materials and Methods). \*, \*\*  $p < 0.05$ ,  $p < 0.01$  by one sample t-test. Error was propagated using standard rules, and error bars represent 95% confidence intervals.

**Fig. 6. Opioid-induced antinociception and respiratory depression.** (A) Antinociception was measured for 180 min using the hot plate assay and increasing doses as noted (in mg/kg XX) of fentanyl, morphine, oliceridine, PZM21, SR-17018, or buprenorphine. Data are normalized to the maximal possible effect (%MPE) and are mean  $\pm$  SEM of three to nine animals (n shown in parentheses for each dosage). (B) Respiratory depression was measured for 240 min using whole-body plethysmography in freely moving animals and increasing doses of fentanyl, morphine, oliceridine, PZM21, SR-17018, or buprenorphine. Data are normalized to baseline respiratory rate and are mean  $\pm$  SEM of five to nine animals (n shown in parentheses). (C) The response of equi-effective doses in antinociception was compared to the respiratory depression assay. \*, \*\*, \*\*\*,  $p < 0.05$ , 0.01, 0.001 by multiplicity-corrected t-test compared to vehicle for peak dose of PZM21, SR-17018, and buprenorphine. All drugs were administered subcutaneously except for SR-17018, which was administered intraperitoneally.

**Fig. 7. Therapeutic windows of opioid ligands.** (A) Dose-response analysis of fentanyl, morphine, oliceridine, PZM21, SR-17018, and buprenorphine in antinociception and respiratory depression assays. Peak effect (Figs. 6A and 6B) is plotted for each dose, with hot-plate (HP) plotted on the left Y axis and whole-body plethysmography (WBP) on the right Y axis. (B) Potency [ $\log(ED_{50})$ ] of each agonist for anti-nociception and respiratory depression. Because curve fit was not possible for SR-17018, potency for antinociception was taken as minimum dose reaching 50% MPE. In the case of respiratory depression induced by buprenorphine and SR-17018, neither of which reached 50% the maximum effect of fentanyl, the highest tested dose is represented as a minimum lower bound of potency.

**Fig. 8. The therapeutic window of opioid ligands is inversely correlated with their intrinsic efficacy.** The therapeutic windows for fentanyl, morphine, oliceridine, PZM21, SR-17018, and buprenorphine were calculated by subtraction of  $\log(ED_{50})$  of antinociception from respiratory depression,  $\Delta\log(ED_{50})$  (Fig. 7) and the error propagated. Therapeutic windows were then plotted against (A) the operational efficacy  $\log(\tau)$  obtained for mini  $G_i$  recruitment or (B) the bias factors,  $\Delta\Delta\log(\tau/K_A)$ . The therapeutic windows of SR-17018 and buprenorphine were estimated and are shown as lower bounds. The relationships of the therapeutic windows of these ligands with the operational efficacies or bias factors at other signaling endpoints is shown in fig. S8. The operational efficacy of fentanyl for  $mG_{si}$  recruitment shown here is an estimated value.

**Fig. 9. Overexpression of GRK2 does not alter bias.** Data from Fig. 1 and Fig. 2 was used to plot the equimolar comparison of the different opioid agonists activities in G protein activation and  $\beta$ -arrestin 2 recruitment in the presence or absence of GRK2.

Commented [AV42]: I revised the legend text to state what was done and what the data shows, not what it means (i.e. that the values correlated) because that is interpretation of the results, which should be limited to the Results section rather than started in the legend.

Commented [MC43R42]: OK

Commented [AV44]: Define the units for the values given in the graph keys.  $\mu\text{M}$ ?

Commented [AV45]: Correct as edited?

Commented [MC46R45]: yes

Commented [AV47]: Correct as edited?

Commented [MC48R47]: yes

**Table 1. Determination of the maximal responses ( $E_{max}$ ) of selected agonists for all the pathways measured at the MOR. Values are expressed as mean  $\pm$  SEM from three to fourteen independent experiments.**

$E_{max}$ % of DAMGO	Nb33	mGsi	GPA	cAMP	GIRK (untreated)	GIRK ( $\beta$ -CNA)
DAMGO	100	100	100	100	100	75 $\pm$ 5
Fentanyl	102 $\pm$ 6	104 $\pm$ 6	103 $\pm$ 2	100 $\pm$ 5	95 $\pm$ 4	74 $\pm$ 7
Methadone	88 $\pm$ 5	107 $\pm$ 3	105 $\pm$ 2	106 $\pm$ 3	ND	ND
Morphine	71 $\pm$ 3	80 $\pm$ 7	98 $\pm$ 5	97 $\pm$ 4	88 $\pm$ 4	56 $\pm$ 3
Oxycodone	70 $\pm$ 4	75 $\pm$ 3	103 $\pm$ 4	106 $\pm$ 7	ND	ND
Oliceridine	42 $\pm$ 8	51 $\pm$ 7	85 $\pm$ 5	86 $\pm$ 5	74 $\pm$ 4	38 $\pm$ 3
PZM21	38 $\pm$ 3	52 $\pm$ 5	86 $\pm$ 7	84 $\pm$ 7	86 $\pm$ 3	40 $\pm$ 2
SR-17018	20 $\pm$ 6	35 $\pm$ 5	61 $\pm$ 13	62 $\pm$ 12	78 $\pm$ 7	33 $\pm$ 6
Buprenorphine	26 $\pm$ 3	36 $\pm$ 3	79 $\pm$ 6	86 $\pm$ 9	54 $\pm$ 5	19 $\pm$ 3

$E_{max}$ % of DAMGO	GRK2	$\beta$ Arr2	Rab5	$\beta$ Arr2 (GRK2)	Rab5 (GRK2)
DAMGO	100	100	100	100	100
Fentanyl	105 $\pm$ 5	70 $\pm$ 4	79 $\pm$ 13	92 $\pm$ 2	88 $\pm$ 4
Methadone	107 $\pm$ 9	88 $\pm$ 3	77 $\pm$ 3	100 $\pm$ 3	98 $\pm$ 6
Morphine	66 $\pm$ 6	28 $\pm$ 3	51 $\pm$ 10	70 $\pm$ 3	75 $\pm$ 5
Oxycodone	64 $\pm$ 8	23 $\pm$ 4	-	72 $\pm$ 5	64 $\pm$ 3
Oliceridine	40 $\pm$ 3	16 $\pm$ 3	-	58 $\pm$ 3	49 $\pm$ 4
PZM21	36 $\pm$ 7	14 $\pm$ 3	-	59 $\pm$ 4	51 $\pm$ 3
SR-17018	41 $\pm$ 9	12 $\pm$ 1	-	49 $\pm$ 7	59 $\pm$ 9
Buprenorphine	28 $\pm$ 5	-	-	34 $\pm$ 2	25 $\pm$ 9

**Table 2. Determination of the potencies (pEC<sub>50</sub>) of selected agonists for all the pathways measured at the MOR. Values are expressed as mean ± SEM from three to fourteen independent experiments.**

pEC <sub>50</sub>	Nb33	mGsi	GPA	cAMP	GIRK (untreated)	GIRK (β-CNA)
<b>DAMGO</b>	6.76 ± 0.04	7.03 ± 0.17	7.94 ± 0.16	8.48 ± 0.29	8.69 ± 0.05	7.29 ± 0.17
<b>Fentanyl</b>	7.36 ± 0.09	7.64 ± 0.10	8.68 ± 0.16	9.20 ± 0.23	9.30 ± 0.09	8.34 ± 0.23
<b>Methadone</b>	6.67 ± 0.05	6.99 ± 0.06	8.00 ± 0.18	8.22 ± 0.17	ND	ND
<b>Morphine</b>	6.67 ± 0.03	6.94 ± 0.05	7.72 ± 0.16	8.10 ± 0.26	7.98 ± 0.10	7.16 ± 0.17
<b>Oxycodone</b>	5.93 ± 0.12	6.22 ± 0.11	6.94 ± 0.23	7.01 ± 0.18	ND	ND
<b>Oliceridine</b>	7.19 ± 0.28	7.46 ± 0.18	8.38 ± 0.19	8.66 ± 0.13	8.43 ± 0.12	7.88 ± 0.33
<b>PZM21</b>	6.88 ± 0.36	7.49 ± 0.14	8.16 ± 0.19	8.64 ± 0.25	8.42 ± 0.07	7.78 ± 0.22
<b>SR-17018</b>	7.48 ± 0.46	7.15 ± 0.33	7.66 ± 0.18	7.67 ± 0.33	6.43 ± 0.13	6.18 ± 0.30
<b>Buprenorphine</b>	8.28 ± 0.21	8.65 ± 0.14	8.85 ± 0.06	9.61 ± 0.37*	7.97 ± 0.21	7.93 ± 0.43

pEC <sub>50</sub>	GRK2	βArr2	Rab5	βArr2 (GRK2)	Rab5 (GRK2)
<b>DAMGO</b>	6.83 ± 0.08	6.27 ± 0.08	6.13 ± 0.15	7.55 ± 0.09	6.97 ± 0.12
<b>Fentanyl</b>	7.33 ± 0.15	7.10 ± 0.06	6.73 ± 0.13	8.08 ± 0.07	7.74 ± 0.07
<b>Methadone</b>	6.76 ± 0.07	6.39 ± 0.11	6.44 ± 0.23	7.15 ± 0.10	7.17 ± 0.27
<b>Morphine</b>	6.78 ± 0.08	6.48 ± 0.27	6.25 ± 0.35	7.31 ± 0.13	6.87 ± 0.10
<b>Oxycodone</b>	6.28 ± 0.09	6.77 ± 0.56	-	6.22 ± 0.06	6.25 ± 0.25
<b>Oliceridine</b>	7.26 ± 0.51	6.45 ± 0.26	-	7.71 ± 0.08	7.21 ± 0.26
<b>PZM21</b>	7.58 ± 0.19	7.36 ± 0.84	-	7.56 ± 0.07	7.56 ± 0.23
<b>SR-17018</b>	6.99 ± 0.19	8.40 ± 0.47	-	6.48 ± 0.46	7.06 ± 0.39
<b>Buprenorphine</b>	7.95 ± 0.28	-	-	8.50 ± 0.21	6.98 ± 6.78



**Table 3. Determination of the efficacies ( $\tau$ ) of selected agonists for all the pathways measured at the MOR.** Values are expressed as mean  $\pm$  SEM from three to fourteen independent experiments. Efficacy of oliceridine, PZM21, SR-17018, and buprenorphine were compared to morphine in a two-way ANOVA with a Holm-Sidak multiple comparisons corrected post-hoc test. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Log $\tau$	Nb33	mGsi	GPA	cAMP	GIRK
Morphine	0.42 $\pm$ 0.08	0.70 $\pm$ 0.07	1.74 $\pm$ 0.22	2.00 $\pm$ 0.31	0.09 $\pm$ 0.05
Oxycodone	0.56 $\pm$ 0.12	0.65 $\pm$ 0.09	1.86 $\pm$ 0.26	1.87 $\pm$ 0.22	ND
Oliceridine	-0.34 $\pm$ 0.12***	-0.12 $\pm$ 0.05**	1.15 $\pm$ 0.25	1.29 $\pm$ 0.23**	-0.24 $\pm$ 0.05
PZM21	-0.33 $\pm$ 0.07***	-0.08 $\pm$ 0.08**	1.18 $\pm$ 0.28	1.44 $\pm$ 0.39*	-0.18 $\pm$ 0.04
SR-17018	-0.86 $\pm$ 0.21***	-0.37 $\pm$ 0.16***	0.66 $\pm$ 0.37***	1.04 $\pm$ 0.28**	-0.28 $\pm$ 0.12
Buprenorphine	-0.62 $\pm$ 0.09***	-0.40 $\pm$ 0.05***	0.63 $\pm$ 0.25***	1.35 $\pm$ 0.39*	-0.61 $\pm$ 0.10**

Log $\tau$	GRK2	$\beta$ Arr2 (GRK2)	Rab5 (GRK2)
Morphine	0.21 $\pm$ 0.07	0.34 $\pm$ 0.07	0.45 $\pm$ 0.18
Oxycodone	0.22 $\pm$ 0.14	0.35 $\pm$ 0.09	0.11 $\pm$ 0.16
Oliceridine	-0.30 $\pm$ 0.18	0.13 $\pm$ 0.03	-0.08 $\pm$ 0.13*
PZM21	-0.37 $\pm$ 0.11	0.11 $\pm$ 0.06	-0.06 $\pm$ 0.18
SR-17018	-0.30 $\pm$ 0.09	-0.04 $\pm$ 0.15	0.04 $\pm$ 0.15
Buprenorphine	-0.57 $\pm$ 0.08	-0.30 $\pm$ 0.09*	-0.84 $\pm$ 0.32***

Commented [AV49]: Table 3 was not called out in the text. Add a callout for it in the Results section.

Commented [MC50R49]: Done

