Low intrinsic efficacy alone cannot explain the improved side effect profiles of new opioid agonists

Edward L. Stahl and Laura M. Bohn

Department of Molecular Medicine, The Scripps Research Institute, 130 Scripps Way, Jupiter, FL, 33458

estahl@scripps.edu; lbohn@scripps.edu

Abstract:

In a recent report in Science Signaling (DOI: 10.1126/scisignal.aaz3140), it was suggested that low intrinsic agonism, and not biased agonism, leads to an improvement in the separation of potency in opioid-induced respiratory suppression versus antinociception. Although many of the compounds that were tested have been shown to display G protein signaling bias in prior publications, the authors conclude that since they cannot detect biased agonism in their cellular signaling studies the compounds are therefore not biased agonists. Rather, they conclude that it is low intrinsic efficacy that leads to the therapeutic window improvement. Intrinsic efficacy is the extent to which an agonist can stimulate a G protein-coupled receptor (GPCR) response in a system, while biased agonism takes into consideration not only intrinsic efficacy, but also potency of an agonist in an assay. Herein, we have re-analyzed the data presented in the published work (DOI: 10.1126/scisignal.aaz3140) (including the recent Erratum: DOI: 10.1126/scisignal.abf9803) to derive intrinsic efficacy and bias factors as $\Delta\Delta\log(\tau/K_A)$ and $\Delta\Delta\log(Emax/EC_{50})$. Based on this reanalysis, the data support the conclusion that biased agonism, favoring G protein signaling, was observed. Moreover, a conservation of rank order intrinsic efficacy was not observed upon comparing responses in each assay, further suggesting that multiple active receptor states were present. These observations agree with prior studies wherein oliceridine, PZM21 and SR-17018 were first described as biased agonists with improvement in antinociception over respiratory suppression in mice. Therefore, the data in the Science Signaling manuscript does provide strong corroborating evidence that G protein signaling bias may be a means to improve opioid analgesia while avoiding certain undesirable side effects.

Introduction

In the March 31, 2020 issue of Science Signaling, a research article by Gillis et al. [1, 2] (hereafter referred to as the SS-manuscript) investigated a series of mu opioid agonists for activity in a compendium of *in vitro* cell-based bioluminescence resonance energy transfer (BRET) studies (summarized in **Table 1**) to evaluate the pharmacological basis of functional selectivity. In particular, the authors focused on two discrete avenues of response: those medicated by G protein signaling or βarrestin2 recruitment. The SS-manuscript also included studies in mice that were intended to compare the therapeutic window of the selected compounds by determining potency in agonist-induced antinociception (hot-plate latency) and respiratory suppression (respiratory frequency) measures. These physiological responses are relevant as they reflect a similar relationship to the therapeutic window observed in clinical use of known opioids. While the study confirmed that the selected biased agonists showed an improved therapeutic window, the authors failed to detect biased agonism in their signaling assays and conclude that intrinsic efficacy and not biased agonism was responsible for the improved therapeutic window [1].

The pursuit of agonists that display G protein signaling over βarrestin2 recruitment was inspired by early observations of mice lacking βarrestin2; where antinociception was enhanced and respiratory suppression was diminished in response to morphine [3-5]. Several laboratories have developed such agonists, including PZM21 [6], a series of Scripps Research "SR" compounds [7, 8], a fungal natural product [9], derivatives of mitragynine [10] and importantly, the first G protein signaling-biased MOR agonist, oliceridine (TRV-130, Olinvyk®,[11]) that has progressed to the clinic [12]. Moreover, an improvement in producing antinociception with fewer signs of respiratory suppression has been seen for these compounds in rodents (with the exception of the fungal agonist which was not tested in vivo). Importantly, with the emerging clinical studies demonstrating an improvement in analgesic efficacy over respiratory suppression for oliceridine [13, 14], it would seem that the early mouse studies were predictive of a useful strategy for improving opioid therapeutics by developing G protein signaling-biased agonists.

Determining whether a compound shows a preference for signaling in one pathway over another has proven to be less straightforward than anticipated. These measures are complicated by the use of overexpression systems in different cellular backgrounds, varying receptor levels, tags and modifications to the receptors and effectors to amplify signal and the assumption that a plethora of assays designed to detect G protein mediated signaling are equal surrogates for that response. This leads to the assumption that the determination of potency (EC_{50}) and efficacy (Emax) across a diverse array of "G protein signaling" assays will define the agonist's ability to stimulate G proteins. In reality, receptor number, signal amplification and the sensitivity of the assay will greatly affect the determination of potency and efficacy; however, if all agonists have an equivalent propensity to activate different pathways, then their rank order efficacy should be preserved (the most potent and efficacious agonist should remain so in all assays) [15-17].

In the SS-manuscript [1], it is proposed that ligand rank order intrinsic efficacy and *not* ligand bias is the driving factor behind the separation of antinociception from respiratory suppression of PZM21, oliceridine, buprenorphine and SR-17018. The authors quantified "biased agonism" and agonist intrinsic efficacy across cellular signaling assays using the operational model originally described by Black and Leff [18]. Notably, the authors reanalyzed efficacy (E_{max}) values from several recent papers (including PZM21 [6] and SR-17018 [8] to support the proposal that it is intrinsic efficacy, and not biased agonism, that leads to improvements in the therapeutic window [1, 19, 20]. *The conclusions of this report [1] and the value of partial agonist efficacy has been the subject of a number of recent review papers [19-21].*

The operational model has become common place in studying agonist performance across a variety of experimental systems and serves as a means to quantify biased agonism. The key feature that makes this model appealing is its capacity to produce a simplified correction for non-linear occupancy-response relationships. The model employs a hyperbolic occupancy-response relationship and assumes that, at a partial level of agonist occupancy, some agonists are able to fully activate their effector proteins. The

product of this analysis are two unique parameters, tau (τ) and K_A. K_A is an equilibrium affinity constant which describes the avidity by which an agonist-receptor complex is formed, while τ is an (intrinsic) efficacy parameter that describes how well the agonist-receptor complex is able to convey a signal to the effector protein. An easy way to understand the scale of τ is that when an agonist has an Emax of 50%, it will have a τ = 1 [22]. This can be realized by substitution of τ = 1 into Equation 1, from Black and Leff [18]:

Equation 1.

$$Emax_{agonist} = Emax_{system} \times \frac{\tau}{1+\tau}, where\ \frac{1}{1+1} = 50\%\ Emax_{system}$$

As the Emax of the agonist approaches 0, the value of tau will approach 0. For most experimental systems, however, τ will be a positive value as inverse agonist receptor reserve is not commonly encountered and would require a different set of fundamental assumptions about the system. Since τ is positive, it can be expressed as log(τ) which will then extend across positive and negative numbers.

In the SS-manuscript, the operational model is employed to calculate intrinsic efficacy (τ) for several partial agonists to assess MOR engagement with G protein or other effectors relative to the 100% maximum system response produced by DAMGO (see Table 1 for assay description, abbreviated names of the signaling assays will be used throughout the text). The resulting values for log(τ) (SS-Table 1), Emax (SS-Table 2), EC_{50} (SS-Table 3), and $\Delta\log(\tau/K_A)$ (SS-Table S1) are presented in the SS-manuscript (for clarity, figures and tables in the original manuscript will be referred to with the prefix: "SS-," i.e. SS-Table 1 [1]). An erratum [2], wherein the following changes were made: SS-Table S1 ($\Delta\log(\tau/K_A)$) values for cAMP for all compounds were replaced, and buprenorphine values were added for the βarr2 and βarr2 (+GRK2); SS-Table 3 (EC_{50}) values for SR-17018 in βarr2 (+GRK2) were changed; no graphs or conclusions were changed in the erratum. All analyses presented herein include the "erratum" values.

Analysis

Discrepancies between τ and Emax

Since the primary conclusion of the manuscript is that intrinsic efficacy, and not agonist bias, correlates with an improved therapeutic window, it was reasonable to investigate these measures of τ and Emax. As noted, the τ value can be calculated from Emax, and vice versa (Equation 1, [1, 18]); therefore, Emax values presented after conversion from SS-Table 1 (conversion using substitution into Equation 1) were compared to the Emax values presented in SS-Table 2 in **Figure 1.** For this comparison, only partial agonists and only the assays that presented all of the values in both SS-Tables were included. While morphine shows the expected consistency between the Emax derived from the two tables; there is considerably less agreement for other agonists.

Interestingly, the Emax is overestimated when derived from the log(τ) and this is consistent for most of the amplified assays (Gαi2 and cAMP) while underestimated for the assays that most proximally detect G protein-receptor interaction (Nb33 and mGsi). The discrepancy is most pronounced for the data provided for SR-17018 where there is significant separation between the Emax prediction from τ (SS-Table 1) and the Emax estimate (SS-Table 2). Furthermore, it is remarkable that certain log(τ) values ± SEM in SS-Table 1 are negative. Specifically, while the Emax for oliceridine was reported as 51% for mGsi in SS-Table 2, a negative log(τ) value of -0.12 ± 0.05 was presented in SS-Table 1 which would require that the Emax be less than 50%, by definition. It should be mentioned that although the mean ± SEM does not overlap zero, this information is secondary to the primary Emax vs. log(τ) comparison. Better stated, a mean Emax > 50% cannot agree with a mean log(τ) < 0 regardless of any consideration of standard error of the mean. Overall, these discrepancies raise questions about the utility of the determination of τ, as presented in SS-Table 1, for assigning rank order intrinsic efficacy.

The use of the operational model centrally assumes a non-linear occupancy-response relationship. As such, an agonist response curve (potency) is expected to reside to the left of the agonist occupancy binding curve (affinity). That is, the agonist affinity is amplified by the system to produce the observed

agonist EC_{50} in a given response. Concern is warranted when a hyperbolic occupancy-response function is employed to analyze what would be better described as a linear occupancy-response relationship i.e., when there is minimal separation between the response and binding curve [23, 20]. Specifically, this can alter the way τ and K_A relate to Emax and EC_{50} due to an inappropriate correction for system amplification. However, as τ and K_A, and Emax and EC_{50}, appear to result from systems with some amplification this concern may not apply. As partial agonists were included in the analysis, an occupancy-response relationship (in a system with some amplification) is likely a close approximation; therefore, the calculations presented herein should be reasonable.

Potency comparisons and interpretations

It is noteworthy that some partial agonists produced EC_{50} values that are orders of magnitude apart (SS-Table 3, **Figure 2**) although in theory these values should be much closer together as partial agonists occupancy-response curves are closely related. It is possible that ligand binding kinetic properties are leading to overestimates of potency in certain assays; this would also artificially affect the perception of bias [24]. The authors demonstrate that some association kinetics were observed to be delayed but dissociation appears to be rapid, as it was readily reversed by naloxone or DAMGO, suggesting that potency was not greatly exaggerated (SS-Figures S2 and S4) [1]. However, it is particularly interesting that naloxone reversal of mGsi was incomplete (compared to baseline) for all agonists tested, with the exception of SR-17018 which was completely reversed by naloxone (SS-Figure S2B). This would suggest that dissociation kinetics may be delayed and lead to overestimates of potency for the other compounds (including the reference agonist) for the mGsi assay, which was subsequently used for bias calculations and comparison to the therapeutic window. Indeed, if this were the case then it is possible that all bias estimates could be significantly underestimated due to lack of reference agonist equilibrium and overestimation of reference agonist potency.

Comparison of efficacy (τ values) derived from provided log(τ), Emax and log(τ/K_A)

For the purposes of bias analysis in the SS-manuscript, the authors focus on five specific responses (Nb33, mGsi, Gαi2 activation, cAMP, and βarrestin2 (+GRK2) recruitment) (see **Table 1** for description of assays). βArrestin2 recruitment is measured in the presence of overexpressed GRK2 as the authors propose that GRK2 overexpression does not alter the analysis of bias (SS-Figure 9). The specific bias analysis that was employed is a reparameterization of the operational model that produces a transduction coefficient, $\log(\tau/K_A)$ [25, 26]. When compared to a reference agonist, this efficacy and affinity integration can be used to normalize and consolidate agonist activity into a single composite parameter, namely the normalized transduction coefficient $\Delta\log(\tau/K_A)$. Further, the analysis used in the SS-manuscript assumes that the responses measured exhibit concentration-response relationships with a Hill slope constrained to one. As such, the authors state that the systems have minimal amplification due to the stochiometric measurement of protein-protein interactions. Constraining the Hill slope to one can be appropriate as $\log(\tau/K_A)$ values have been shown to remain linear across a wide range of Hill slope values including one [26]. Of note, upon producing simulated concentration response curves to reflect the data in the SS-manuscript, the Hill slopes do not appear to be broadly variable (**Supplemental Figures 1 and 2**).

When the normalized transduction coefficient was determined in the SS-manuscript, an estimate for agonist affinity (K_A) is also produced. It can be seen that with $\mathrm{Log}(\tau/K_A)$ and K_A known, τ and $\log(\tau)$ can subsequently be determined by Equations 2 and 3:

Equation 2.

$$\tau = 10^{log\left(\frac{\tau}{K_A}\right)} \times K_A$$

Equation 3.

$$\log(\tau) = log\left(10^{log\left(\frac{\tau}{K_A}\right)} \times K_A\right)$$

A recent review article highlighted the relationship between the three (or four)-parameter equation and the values that result from the operational analysis [20]. This follows from the parameter definition originally described by Black and Leff, [18] in Equation 1 (shown above) and Equation 4:

Equation 4.

$$EC_{50} = \frac{K_A}{1+\tau}$$

It is possible to estimate K_A using the EC_{50} values from SS-Table 3 by applying Equation 5:

Equation 5.

$$K_A = EC_{50} \times (1+\tau)$$

More interestingly, since $\log(\tau/K_A)$ was produced for a number of agonists in SS-Table S1, it is possible to solve for τ directly by using Equation 6:

Equation 6.

$$\tau = 10^{log\left(\frac{\tau}{K_A}\right)} \times \left(EC_{50} \times (1+\tau)\right)$$

It should be pointed out that these equations are most appropriate for partial agonists as the Emax of a partial agonist is defined at 100% occupancy. Therefore, τ was only estimated for agonists that produced less than 90% efficacy wherein the maximum response could be assumed to be at levels that saturate the available receptor population. Using the $\log(\tau/K_A)$ for the partial agonists (SS-Table S1), the Emax (SS-Table 2), and the pEC_{50} (SS-Table 3), the τ parameter can be independently determined; a comparison of these derivations of τ are shown in **Figure 3**. Immediately, it is evident that there is poor agreement between τ values derived from the three sources for the Gαi2 and cAMP measures and this was least consistent for SR-17018.

As the focus of the SS-manuscript was the correlation of agonist efficacy or bias with therapeutic window, it was reasonable to produce the $\Delta\log(\tau/K_A)$ values that are used to evaluate functional selectivity (bias). As such, an estimate of $\log(\tau/K_A)$ can be derived from the three-parameter curve fit using Equation 7:

Equation 7.

$$\log\left(\frac{\tau}{K_A}\right) = log\left(\frac{\frac{\frac{Emax_{agonist}}{Emax_{system}}}{1-\frac{Emax_{agonist}}{Emax_{system}}}}{EC_{50} \times \left(1 + \frac{\frac{Emax_{agonist}}{Emax_{system}}}{1 - \frac{Emax_{agonist}}{Emax_{system}}}\right)}\right)$$

This equation is easily adapted to Microsoft® Excel for both parameter conversion and the further bias analysis that is the principal endpoint. For completeness, all Microsoft® Excel spreadsheets for all figures used for the analyses, herein, are provided for examination in the supplement.

Comparison of $\Delta\log(\tau/K_A)$ derived from EC_{50} and Emax compared to $\Delta\log(\tau/K_A)$ values presented

In bias analysis, the $\Delta\log(\tau/K_A)$ values specifically provide a discrete measure of each agonists activity in each experimental system. The Δ in the name comes from the correction to the activity of the reference agonist (usually a full agonist where τ cannot be explicitly determined) within each system. The $\log(\tau/K_A)_{Reference}$ is defined as the pEC_{50} of the reference agonist in each system and the $\Delta\log(\tau/K_A)$ for each test agonist is expressed in Equation 8 as:

Equation 8.

$$\Delta Log\left(\frac{\tau}{K_A}\right) = Log\left(\frac{\tau}{K_A}\right)_{Test} - Log\left(\frac{\tau}{K_A}\right)_{Reference}$$

In **Figure 4,** the $\Delta\log(\tau/K_A)$ for each agonist in each response is plotted on the X-axis (calculated from SS-Table S1). In cases where $\Delta\log(\tau/K_A)$ was determined for full agonists, the comparison is presented as $\Delta pEC_{50(Test\text{-}Reference)}$. The Y-axis presents the $\Delta\log(\tau/K_A)$ that was produced using equations 7 and 8 with EC_{50}, $Emax_{system}$, and Emax values from SS-Tables 2 and 3. The $Emax_{system}$ is defined as the response produced by DAMGO in each system (100%). The coefficient of determination, (R^2), when the slope is constrained to one, is also presented in each panel to demonstrate the strength of the agreement between $\Delta\log(\tau/K_A)$ presented in SS-Table S1 and $\Delta\log(\tau/K_A)$ derived from the SS-Tables 2 & 3. Notably, the

values do not agree for the Gai2 (R^2= 0.3638) and cAMP (R^2=0.7182) measures; values for SR-17018 are in disagreement in most assays.

Comparison of $\Delta\Delta\log(\tau/K_A)$ presented in tables, graphs and derived from Emax and EC_{50} values

Having normalized the performance of each agonist in each assay to the performance of the reference agonist DAMGO (mean $\Delta\log(\tau/K_A)$), it is further possible to compare an agonist's relative performance between two assays using Equation 9:

Equation 9.

$$\Delta\Delta log\left(\frac{\tau}{K_A}\right)_{assay1-assay2} = \Delta log\left(\frac{\tau}{K_A}\right)_{assay1} - \Delta log\left(\frac{\tau}{K_A}\right)_{assay2}$$

As written in Equation 9, a positive number would indicate a preference for *assay 1* and a negative number would indicate a preference for *assay 2*. When the $\Delta\Delta\log(\tau/K_A)$ and its 95% confidence intervals do not include zero the value is indicative of an agonist that exhibits selectivity between one response and another (i.e. functional selectivity, biased agonism) [26]. More broadly stated, $\Delta\Delta\log(\tau/K_A) \neq 0$ indicates that the agonist promotes a discrete selection of an active state that prefers one effector over another; as such $\Delta\Delta\log(\tau/K_A)$ is used to quantify "biased agonism." The $\Delta\Delta\log(\tau/K_A)$ is frequently expressed as extending toward or away from one response or another, depending on perspective of preferring or not preferring a receptor active state. This concept was first proposed mechanistically by a number of studies both in brain and recombinant cell systems [17], [27].

Therefore, using the process described above, the bias factor (as $\Delta\Delta\log(\tau/K_A)$), can be calculated from the data included in SS-Tables 2, 3 and S1. First, using SS-Table S1, the $\Delta\log(\tau/K_A)$ value of βarr2+GRK2 can be subtracted from the $\Delta\log(\tau/K_A)$ for each of the G protein signaling pathways to generate the $\Delta\Delta\log(\tau/K_A)$ (**Equation 9**) to determine preference over βarr2+GRK2 recruitment. These calculated parameters ($\Delta\Delta\log(\tau/K_A)$ are presented in **Figure 5A** and are compared to values graphed in SS-Figure 5 (as estimated from the graphic because the $\Delta\Delta\log(\tau/K_A)$ values were not included in the SS-

manuscript). Buprenorphine was not included in this analysis as a $\log(\tau/K_A)$ value for buprenorphine in the βarr2 assays was not provided in SS-Table S1. Although the error of these calculations is not readily accessible, the error propagation does not directly inform or affect the means comparison, as discussed above.

Given the discrepancy between calculated $\Delta\Delta\log(\tau/K_A)$ values and those graphed in SS-Figure 5, and a concern that perhaps the SS-Table S1 data were incorrectly derived in the published paper, $\Delta\log(\tau/K_A)$ values were then calculated from the Emax and EC_{50} values presented in SS-Tables 2 & 3 as described in **Equation 7**. The bias factors ($\Delta\Delta\log(\tau/K_A)$) were again produced by simple subtraction (**Equations 8 & 9**) and are presented **Figure 5B**. The calculations that produced all values determined in **Figure 5** are provided in the supplemental Microsoft® Excel files.

The comparison of $\Delta\Delta\log(\tau/K_A)$ (**Figure 5**) taken from the three different sources in the manuscript (SS-Figure 5 (simulation), SS-Table S1 (subtraction), and SS-Tables 2&3 (**Equation 7** using Emax and EC_{50}), reveal that values are fairly consistent but with some striking exceptions. The values presented in SS-Figure 5 consistently underestimate the G protein signaling preference for SR-17018 across all assays compared to the data obtained from SS-Table S1 (**Figures 5A & B**). The SS-manuscript concludes that SR-17018 "showed no statistically significant bias toward or away from any G protein activation measure." If the same error in SS-Figure 5 were applied to the newly calculated $\Delta\Delta\log(\tau/K_A)$, the 95% confidence interval (which spans more than two orders of magnitude) would overlap zero even though the new mean calculated $\Delta\Delta\log(\tau/K_A)$ changes substantially (by an order of magnitude, in some cases). As the 95% confidence intervals for $\Delta\Delta\log(\tau/K_A)$ values are not symmetrical (SS-Figures 5, S6), it is not clearly described how these intervals were determined. However, the standard method proposed[26] was not used, indicating that a direct reassignment to these newly calculated mean values would not be possible or useful.

This would lead to the general conclusion that the presences of functional selectivity cannot be proven or disproven, based solely on this analysis, as the error is so large as to preclude *any* meaningful interpretation.

Comparison of efficacy and therapeutic window

The primary goal of the SS-manuscript was to demonstrate that intrinsic efficacy, and not bias, was responsible for the improvement in the therapeutic window of the biased MOR agonists. To that end, the manuscript produced a series of in vivo experiments to determine therapeutic window as measured by the $\Delta LogED_{50}$ between antinociception (hot plate latency) and respiratory suppression (respiratory frequency). These therapeutic window values were then correlated with ligand efficacy, log(τ), as well as functional selectivity (bias, $\Delta\Delta\log(\tau/K_A)$), between different responses.

For this comparison, the authors focused on intrinsic efficacy in the mGsi assays and bias comparison between mGsi and βarr2+GRK2 assays in SS-Figure 8. **Figure 6** presents a representation of the intrinsic efficacy, log(τ), taken from SS-Table 1 or derived from SS-Table 2 (Emax) graphed for mGsi (**Figure 6A**) and βarr2+GRK2 (**Figure 6B**); this is compared with therapeutic window (**Figure 6C**). It should be noted that it was not possible to derive the log(τ) value for fentanyl in the mGsi assay as the maximum response of fentanyl intersects with the maximal response of the system. As such the log(τ) value for fentanyl, in assays where it approaches or exceeds 100%, will approach infinity. The Emax for the βarr2+GRK2 assay (Emax= 92% in SS-Table 2) was used to estimate log(τ) in **Figure 6B**; no value was provided in SS-Table 1. As the authors point out in SS-Figure 3 of the manuscript, the partial agonists are partial in both G protein signaling assays as well as in βarrestin2 recruitment. It can therefore be interpreted that decreasing the intrinsic efficacy for recruiting βarrestin2 as well as G protein correlates with an improved therapeutic window (**Figure 6A-C**). However, in both cases, there is not much refinement in the assessment, as PZM21 shows a wider therapeutic window, yet the intrinsic efficacy is the same as that observed for oliceridine in both assays.

Since therapeutic window values were not included in the manuscript, the plot in **Figure 6C** was created by tracing and estimating values that correspond with SS-Figure 8B; error is also included based on this same method. Additionally, the mean therapeutic window values for SR-17108 and buprenorphine were not included as an estimate of ΔED_{50} could not be calculated (as these agonists did not produce 50% of maximum respiratory suppression). Therefore, it is not possible to estimate either the mean or the error for the ΔED_{50} for both SR-17018 or buprenorphine as their means approach infinity. The therapeutic window for these compounds was not measured; any representation of therapeutic windows that does define either of these parameters is, at the minimum, an overinterpretation of the data. With this recognition, the therapeutic window graph in **Figure 6C** presents the lower confidence intervals for means that extend to infinity as depicted in SS-Figure 8.

Figure 6D and 6E are from **Figure 5** and are plotted, as presented in SS-Figure 8, for comparison to the therapeutic window plot (shown again within the figure). **Figure 6D** shows the $\Delta\log(\tau/K_A)$ values calculated from SS-Table S1 ($\log(\tau/K_A)$) and these are shown in comparison to the values that were presented in graphical form in SS-Figures 5 and 8B. Here it can be clearly seen, using the Emax and EC_{50} values to derive $\log(\tau/K_A)$, that the biased ligands show an improvement in therapeutic window, moreover, the difference in bias between PZM21 and oliceridine now reflect the differences observed in therapeutic window.

Comparison of relative intrinsic activity and therapeutic window

In this reanalysis of the data presented in the SS-manuscript, efforts were made to rederive τ and K_A from E_{max} and EC_{50}, presented in the paper, to arrive at the $\Delta\Delta\log(\tau/K_A)$ bias factors (**Figure 5**). However, given that these are partial agonists, it is also possible to use the values directly produced by the three-parameter concentration-response function to determine relative intrinsic activity [25] in Equation 10:

Equation 10.

$$\Delta\Delta log\left(\frac{E_{max}}{EC_{50}}\right)_{assay1-assay2} = \Delta log\left(\frac{E_{max}}{EC_{50}}\right)_{assay1} - \Delta log\left(\frac{E_{max}}{EC_{50}}\right)_{assay2}$$

Relative intrinsic activity provides a means of interpreting the activity of agonist in different systems without necessarily fully characterizing the stimulus-response relationship. As such, it is an empirical method of identifying agonists that show preference for a subset of responses compared to other agonists that remain impartial between different responses. The results of this analysis can be visualized in **Figure 7A** comparing $\Delta\Delta\log(E_{max}/EC_{50})$**,** where a preference for G protein signaling over βarr2 (+GRK) is observed across the four platforms used to assess G protein signaling. This analysis is in direct agreement with the $\Delta\Delta\log(\tau/K_A)$ values derived from analysis of E_{max} and EC_{50}, to generate τ and K_A in **Figure 5B**.

The authors chose to focus on comparing responses to βarr2 only in the presence of GRK2. GRK2 overexpression is included in the βarr2 assay to improve system sensitivity, as GRK2 phosphorylates MOR and facilitates βarr2 recruitment [1, 28]. Therefore, it is expected that while GRK2 overexpression will enhance agonist potency and efficacy, the degree of enhancement should be *consistent* between the agonists. Preference for βarr2 recruitment in the absence of GRK, compared to GRK overexpression, is presented in **Figure 7B.** Not only is there divergence from unity ($\Delta\Delta\log(E_{max}/EC_{50}) \neq 0$); it is readily apparent that a number of agonists exhibit a preference for βarr2 recruitment in the *absence* of GRK2 relative to their activity in the presence of GRK2 overexpression. This would suggest that GRK2 overexpression could be acting to *decrease* system sensitivity and diminish the βarr2 recruitment produced by the agonist, which is in direct contrast to the expected role of GRK2 in this system (to facilitate βarr2 recruitment). Overall, it is unusual that GRK overexpression worsens an agonist's ability to recruit βarr2 to the receptor.

We also note that most of the agonists show a preference away from GIRK recruitment and towards βarr2 +GRK2, therefore we asked whether the agonists would demonstrate functional selectivity between the G protein signaling assays. While GIRK was considered an orthogonal assay for G protein-mediated

signaling, the data demonstrate robust signaling bias between the ability to activate the other G protein pathways over GIRK (**Figure 7C**). This is particularly intriguing as activation of GIRK has been implicated in MOR-mediated respiratory suppression [29]; indeed, as presented in **Figure 7D**, the bias away from GIRK activation aligns well with the therapeutic window spectra presented in **Figure 6**.

Active state affinity and rank order efficacy are not preserved.

Because the conclusions of the SS manuscript are that intrinsic activity is directly related to therapeutic index, it was relevant to directly compare intrinsic efficacy observed across the different assays. Moreover, since most of the ligands of interest generally act as partial agonists, it was reasonable to generate the active state selectivity predictions that would result from analysis with the two-state model [17]. The active state selectivity ratio (α) can be calculated [15] as shown in Equation 11:

Equation 11.

$$\alpha = \frac{K^*}{K} = \frac{\frac{E_{Max}}{E} - 1}{L}$$

where K* is the active state affinity and K is the inactive state affinity. E_{max} is the maximum response of the full agonist and E is the maximum response of the agonist of interest (from SS-Table 2) and L is the system intrinsic isomerization ratio (R/R*) and was estimated to be 100 for all responses [15, 17]. Further, with the calculated active state selectivity ratio (α), it is possible to use the EC_{50} to calculate the active state affinity constant of the agonist for the receptor as shown in Equation 12 [17]:

Equation 12.

$$K^* = \frac{EC_{50} \cdot \left(1 + L\frac{K^*}{K}\right)}{1 + L}$$

where K* is the agonist affinity constant for the active state in each response using the EC_{50} that is produced from each response. The K*/K is calculated in Equation 11 where, again, L is set equal to 100.

Conceptually, the potency and efficacy of full agonists are highly sensitive to receptor reserve and system sensitivity. That is, a small number of receptors saturated with a full agonist can lead to a full response of the system at a concentration that is less than the active state affinity constant. Partial agonists are less subject to this type of system dependent amplification as their activity is directly proportional to their occupancy. For this reason, it would be expected that the active state affinity (K*) of a partial agonist would be generally conserved across multiple assays, despite differences in signal amplification. However, **Figure 8A** shows that the active state affinity (K*) of several of the partial agonists studied spans two orders of magnitude. This suggests that the differences in affinity may be due to multiple active states of the receptor as previously proposed [17].

However, even if partial agonists could exhibit different affinities for the active state of the receptor (due to system-dependent parameters including receptor-reserve or threshold sensitivity), *the rank order of agonist efficacy should remain constant in different systems if a single active state receptor is responsible for mediating all responses* [16, 17]. **Figures 8B-C** present the comparison of efficacy as E_{max} and τ (derived from SS-Table S1) where it is clear that the relative intrinsic activity (also known as rank order efficacy) is not conserved across the assays. It should be noted that estimates of intrinsic efficacy are not subject to agonist binding kinetics as these values are determined at saturation. These findings distinctly and explicitly implicate more than one active state of the receptor being responsible for stimulating different responses; moreover, the differences in active state selectivity are directly related to the distinction of biased agonists.

Discussion:

The comparison of agonist-induced activity across multiple assays to determine functional selectivity can be a daunting task. While many assays are developed to enhance signal to noise ratios, this can come at a cost of signal amplification. Moreover, certain assays are not amenable to developing stable

cells lines and necessitate repeated transient transfections, which can in turn alter receptor number and assay sensitivity. Therefore, it is imperative that comparisons are always made to a reference agonist's performance (or more than one reference agonist) and that assays are performed in parallel (to avoid changes in the systems over time). The most important requirement for determining functional selectivity is actually two-fold: high quality consistent data and the appropriate application of mechanistic pharmacology analysis. In this evaluation of existing data, we have presented several known approaches for analyzing data to identify a divergence in receptor signaling across assays and have presented an additional analysis adapted from the two-state model of receptor activation (K*). The benefit of determining K* is that is produces a microaffinity constant for the active state preference over the inactive state of the receptor. This allows for normalization of efficacy and potency to produce affinity estimates. If there is no functional selectivity, it would be expected that the rank order of these affinity estimates would be conserved across all systems that are being investigated. Divergence from this conservation will indicate a preference for an assay-dependent active state. Further, the utilization of mechanistic analysis to underpin more empirical analysis can lend a degree of confidence to the general conclusion that ligand bias is observed within a set of data.

In the re-calculation of bias from the SS-manuscript it was not possible to interpret error for the data set as there was only limited information provided by the authors; for example, the number of experiments were provided as a range (3-14 replications) and the S.E.M. was presented in the tables. However, since the derivation of τ and K_A is dependent on the potency (EC_{50}) and efficacy (Emax) within each assay the ability to produce a "bias factor" with any reasonable error of certainty will require that the individual assays are: first, reproducible, and secondly, adequately powered by replicates. There is reasonable concern that the experimental measures presented herein exhibit large error in the estimation of potency (pEC_{50} in SS-Table 3) and efficacy (Emax, SS-Table 2). This can be visualized in **Figure 2**, where the error in potency exceeds a half of log order in several assays. Since the authors report running 3-14 experiments per data set, it would seem that perhaps, in cases where error is very high, that the studies were

underpowered. This initial error propagates throughout additional analyses, leading to 95% CI that are greater than the mean (SS-Figure 5, SS-Figure S6). Therefore, it is apparent that several assays reported in the SS-manuscript are underpowered and do not support the conclusion that there is no "significant" bias observed for any of the compounds. This concern with reproducibility is further supported by the degree of error present upon averaging the parameters derived from DAMGO response curves. DAMGO, as the reference agonist, would be included in every experiment (presumably accounting for the upper n=14 described in SS-Tables 1-3); yet, the margin of error in DAMGO's potency was wider than expected for several of the signaling assays **(Figure 2,** SS-Table 3**)**. The reproducibility issues may stem from the use of transient transfection systems with variable receptor density and effector expression levels. Moreover, there was no indication that protein and receptor expression levels were monitored for consistency between experiments.

In many ways, the SS-manuscript reproduces observations from several investigators that have proposed that improving G protein signaling while avoiding βarrestin2 recruitment may be a useful approach to limit respiratory suppression while preserving antinociception [3-6, 8, 11]. Specifically, the authors reproduce in vivo observations originally published for SR-17018 [8], oliceridine [11], and PZM21 [6]. Moreover, the demonstration that PZM21, at doses up to 100 mg/kg, produces less than half of the respiratory suppression induced by 10 mg/kg morphine; this is very encouraging as it had only been tested up to 40 mg/kg in the first study, where it showed no respiratory suppression compared to vehicle [6]. It should be noted that these observations are in direct contrast to a study that showed PZM21 and morphine producing equivalent respiratory suppression at 10 mg/kg [30].

The idea that decreasing intrinsic efficacy may serve to improve the therapeutic utility of mu opioid receptor agonist to manage pain and limit side effects has been pursued for some time, resulting in compounds such as buprenorphine [31]. However, caution should be maintained in translating efficacy

observed in cultured systems to expectations observed in clinical use. In particular, buprenorphine is fully efficacious in the treatment of some pain conditions [32], as is morphine, even though both perform as partial agonists across many assay systems, as demonstrated in the SS-manuscript. As such, there has not been a direct correlation between partial agonism observed in different cell-based assays and the effects of a drug *in vivo*; therefore, it may not be fully predictive that a partial agonist may be safer regarding respiratory suppression. Indeed, buprenorphine produces respiratory suppression in humans [33], with the additional complication that it is difficult to reverse with naloxone [34] as buprenorphine is a long lasting, high affinity agonist [31]. In rats, buprenorphine produces respiratory suppression to the same extent as fentanyl which would also seem to contradict the efficacy hypothesis as well as the biased agonism hypothesis [35]. However, studies with buprenorphine are further complicated by the fact that it has activity at other opioid receptors and also has nonselective active metabolites that can suppress respiration [36-38].

The conclusion of the SS-manuscript is that partial agonism is responsible for the avoidance of respiratory suppression, and this is supported by the choice of agonists and how those agonists performed in the G protein signaling assays selected. However, the same could be said for the efficacy in the βarrestin2 recruitment assays, as the biased agonists act as partial agonists there as well, in a manner that also aligns with an improved therapeutic window. The data presented here may support the role of agonist efficacy in determining the extent of respiratory suppression in mice; however, no counter argument was purposefully explored (i.e. testing of a full agonist with less respiratory suppression or a partial agonist with more respiratory suppression).

The study by Schmid et al., 2017, addressed this question by generating a series of structurally related MOR agonists that spanned a spectrum of both bias and intrinsic efficacy [8]. One example from that study is SR-11501, a partial agonist that has similar potency and efficacy to morphine in G protein signaling and in antinociception studies. However, SR-11501 displays bias towards βarrestin2 recruitment over

G protein signaling and produces fentanyl-like respiratory suppression, resulting in a very narrow therapeutic window. Moreover, examples of two full agonists that display bias towards G protein signaling, SR-14968 and SR-14969, were shown to produce a wider therapeutic window than morphine (less respiratory suppression vs. antinociception) [8]. Schmid et al. also showed that fentanyl, which has a narrow therapeutic window, acts as a partial agonist in mouse brainstem membranes using ^{35}S-GTPγS binding assays [8]. Fentanyl has also been shown to act as a partial agonist in spinal cord [39] and in rat thalamus [40] in studies assessing ^{35}S-GTPγS binding. Therefore, intrinsic efficacy can be uncoupled from biased agonism and, within the SR series of agonists, G protein signaling bias was shown to correlate with decreased respiratory suppression, independent of intrinsic efficacy [8].

Regardless of the distinction of bias, the comparison of active state affinity constants and the observation that rank order efficacy was not conserved across the spectrum of assays investigated provides compelling support for the idea that these experimental compounds present interesting and substantially novel pharmacology. Ultimately, the determination of biased agonism, or intrinsic efficacy for that matter, will be dependent on the cellular signaling systems used (i.e. receptor density, effector expression, amplification of signal) as well as reference agonist used to determine the maximum possible signaling response in the system. It is not unexpected that perceptions of bias and efficacy will vary between studies [22, 41]. In a recent example, similar BRET-based studies were elegantly employed for similar analysis by Ehrlich et al., [42] investigating ligand bias at the mu opioid receptor where they demonstrate diverse signaling profiles of oliceridine, PZM21 and buprenorphine in neurons. In time, the value of these cellular readouts as predictors of agonist function may be supported by detecting state-dependent structural conformations of MOR induced by biased ligands binding to the receptor [43-45]. This does not make their pharmacology less remarkable not does it disprove the hypothesis that ligands may demonstrate active state selectivity. On the contrary, as more studies present interesting and novel findings regarding the spectrum of activity these compounds demonstrate, it may be possible to more definitively establish these differences.

In particular, the demonstration of changes in rank order efficacy is a compelling finding that further substantiates and validates the value of these compounds. From a practical standpoint, however, the correlation between the pharmacological properties and the improvement in therapeutic efficacy while limiting side effects in human studies should drive further investigations. Indeed, oliceridine, which is a partial agonist in both G protein signaling and βarrestin2 recruitment, as well as a biased agonist for G protein signaling, is producing superior analgesia/respiratory suppression in post-operative pain [11, 13, 14]. Therefore, while partial agonism could be a desirable pharmacological property, biased agonism may serve as a means to further improve the therapeutic window.

Table 1. Assay abbreviations from the SS-manuscript [1].

ASSAY	MEASURE	EFFECTOR	MOR	CELLS	INDICATOR
Nb33	Giα nanobody recruitment	Nb33-venus	flag-mMOR-NLuc	transient HEK293	BRET
mGsi	mini soluble Gi protein recruitment	NES-venus-mGsi	flag-mMOR-NLuc	transient HEK293	BRET
Gai2	association of Gβ with GRK3ct and dissociation from Gαi2	Gαi2 venus-155-239-Gβ1 venus-1-155-Gγ2 and	flag-mMOR-NLuc	transient HEK293	BRET
cAMP	cAMP biosensor	CAMYEL YFP-Epac-Rluc	flag-mMOR	transient HEK293	BRET
βarr2+GRK2	βarrestin2 recruitment with overexpressed GRK2	βarr2-YFP (bovine)	flag-MOR-NLuc	transient HEK293	BRET

Funding. This work is funded by NIH, NIDA grants: DA033073 and DA038964.

References.

[1] Gillis, A., Gondin, A. B., Kliewer, A., Sanchez, J., Lim, H. D., Alamein, C., Manandhar, P., Santiago, M., Fritzwanker, S., Schmiedel, F., Katte, T. A., Reekie, T., Grimsey, N. L., Kassiou, M., Kellam, B., Krasel, C., Halls, M. L., Connor, M., Lane, J. R., Schulz, S., Christie, M. J., and Canals, M. (2020) Low intrinsic efficacy for G protein activation can explain the improved side effect profiles of new opioid agonists, *Sci Signal 13*.

[2] Gillis, A., Gondin, A. B., Kliewer, A., Sanchez, J., Lim, H. D., Alamein, C., Manandhar, P., Santiago, M., Fritzwanker, S., Schmiedel, F., Katte, T. A., Reekie, T., Grimsey, N. L., Kassiou, M., Kellam, B., Krasel, C., Halls, M. L., Connor, M., Lane, J. R., Schulz, S., Christie, M. J., and Canals, M. (2021) Erratum for the Research Article: "Low intrinsic efficacy for G protein activation can explain the improved side effect profiles of new opioid agonists" *Science Signaling 14*, eabf9803.

[3] Bohn, L. M., Lefkowitz, R. J., Gainetdinov, R. R., Peppel, K., Caron, M. G., and Lin, F. T. (1999) Enhanced morphine analgesia in mice lacking beta-arrestin 2, *Science 286*, 2495-2498.

[4] Bohn, L. M., Gainetdinov, R. R., Lin, F. T., Lefkowitz, R. J., and Caron, M. G. (2000) Mu-opioid receptor desensitization by beta-arrestin-2 determines morphine tolerance but not dependence, *Nature 408*, 720-723.

[5] Raehal, K. M., Walker, J. K., and Bohn, L. M. (2005) Morphine side effects in beta-arrestin 2 knockout mice, *J Pharmacol Exp Ther 314*, 1195-1201.

[6] Manglik, A., Lin, H., Aryal, D. K., McCorvy, J. D., Dengler, D., Corder, G., Levit, A., Kling, R. C., Bernat, V., Hubner, H., Huang, X. P., Sassano, M. F., Giguere, P. M., Lober, S., Da, D., Scherrer, G., Kobilka, B. K., Gmeiner, P., Roth, B. L., and Shoichet, B. K. (2016) Structure-based discovery of opioid analgesics with reduced side effects, *Nature 537*, 185-190.

[7] Kennedy, N. M., Schmid, C. L., Ross, N. C., Lovell, K. M., Yue, Z., Chen, Y. T., Cameron, M. D., Bohn, L. M., and Bannister, T. D. (2018) Optimization of a Series of Mu Opioid Receptor (MOR) Agonists with High G Protein Signaling Bias, *J Med Chem 61*, 8895-8907.

[8] Schmid, C. L., Kennedy, N. M., Ross, N. C., Lovell, K. M., Yue, Z., Morgenweck, J., Cameron, M. D., Bannister, T. D., and Bohn, L. M. (2017) Bias Factor and Therapeutic Window Correlate to Predict Safer Opioid Analgesics, *Cell 171*, 1165-1175 e1113.

[9] Dekan, Z., Sianati, S., Yousuf, A., Sutcliffe, K. J., Gillis, A., Mallet, C., Singh, P., Jin, A. H., Wang, A. M., Mohammadi, S. A., Stewart, M., Ratnayake, R., Fontaine, F., Lacey, E., Piggott, A. M., Du, Y. P., Canals, M., Sessions, R. B., Kelly, E., Capon, R. J., Alewood, P. F., and Christie, M. J. (2019) A tetrapeptide class of biased analgesics from an Australian fungus targets the micro-opioid receptor, *Proc Natl Acad Sci U S A 116*, 22353-22358.

[10] Varadi, A., Marrone, G. F., Palmer, T. C., Narayan, A., Szabo, M. R., Le Rouzic, V., Grinnell, S. G., Subrath, J. J., Warner, E., Kalra, S., Hunkele, A., Pagirsky, J., Eans, S. O., Medina, J. M., Xu, J., Pan, Y. X., Borics, A., Pasternak, G. W., McLaughlin, J. P., and Majumdar, S. (2016) Mitragynine/Corynantheidine Pseudoindoxyls As Opioid Analgesics with Mu Agonism and Delta Antagonism, Which Do Not Recruit beta-Arrestin-2, *J Med Chem 59*, 8381-8397.

[11] DeWire, S. M., Yamashita, D. S., Rominger, D. H., Liu, G., Cowan, C. L., Graczyk, T. M., Chen, X. T., Pitis, P. M., Gotchev, D., Yuan, C., Koblish, M., Lark, M. W., and Violin, J. D. (2013) 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.

[12] Lambert, D., and Calo, G. (2020) Approval of oliceridine (TRV130) for intravenous use in moderate to severe pain in adults, *Br J Anaesth*.

[13] Ayad, S., Demitrack, M. A., Burt, D. A., Michalsky, C., Wase, L., Fossler, M. J., and Khanna, A. K. (2020) Evaluating the Incidence of Opioid-Induced Respiratory Depression Associated with

Oliceridine and Morphine as Measured by the Frequency and Average Cumulative Duration of Dosing Interruption in Patients Treated for Acute Postoperative Pain, *Clin Drug Investig 40*, 755-764.
[14] Bergese, S., Berkowitz, R., Rider, P., Ladouceur, M., Griffith, S., Segura Vasi, A., Cochrane, K., Wase, L., Demitrack, M. A., and Habib, A. S. (2020) Low Incidence of Postoperative Respiratory Depression with Oliceridine Compared to Morphine: A Retrospective Chart Analysis, *Pain Res Manag 2020*, 7492865.
[15] Leff, P. (1995) The two-state model of receptor activation, *Trends Pharmacol Sci 16*, 89-97.
[16] Leff, P., Scaramellini, C., Law, C., and McKechnie, K. (1997) A three-state receptor model of agonist action, *Trends Pharmacol Sci 18*, 355-362.
[17] Berg, K. A., Maayani, S., Goldfarb, J., Scaramellini, C., Leff, P., and Clarke, W. P. (1998) Effector pathway-dependent relative efficacy at serotonin type 2A and 2C receptors: evidence for agonist-directed trafficking of receptor stimulus, *Mol Pharmacol 54*, 94-104.
[18] Black, J. W., and Leff, P. (1983) Operational models of pharmacological agonism, *Proc R Soc Lond B Biol Sci 220*, 141-162.
[19] Gillis, A., Kliewer, A., Kelly, E., Henderson, G., Christie, M. J., Schulz, S., and Canals, M. (2020) Critical Assessment of G Protein-Biased Agonism at the mu-Opioid Receptor, *Trends Pharmacol Sci*.
[20] Gillis, A., Sreenivasan, V., and Christie, M. J. (2020) Intrinsic Efficacy of Opioid Ligands and Its Importance for Apparent Bias, Operational Analysis, and Therapeutic Window, *Mol Pharmacol 98*, 410-424.
[21] Faouzi, A., Varga, B. R., and Majumdar, S. (2020) Biased Opioid Ligands, *Molecules 25*.
[22] Kenakin, T., and Christopoulos, A. (2013) Signalling bias in new drug discovery: detection, quantification and therapeutic impact, *Nat Rev Drug Discov 12*, 205-216.
[23] Kelly, E. (2013) Efficacy and ligand bias at the mu-opioid receptor, *Br J Pharmacol 169*, 1430-1446.
[24] Klein Herenbrink, C., Sykes, D. A., Donthamsetti, P., Canals, M., Coudrat, T., Shonberg, J., Scammells, P. J., Capuano, B., Sexton, P. M., Charlton, S. J., Javitch, J. A., Christopoulos, A., and Lane, J. R. (2016) The role of kinetic context in apparent biased agonism at GPCRs, *Nat Commun 7*, 10842.
[25] Griffin, M. T., Figueroa, K. W., Liller, S., and Ehlert, F. J. (2007) Estimation of agonist activity at G protein-coupled receptors: analysis of M2 muscarinic receptor signaling through Gi/o,Gs, and G15, *J Pharmacol Exp Ther 321*, 1193-1207.
[26] Kenakin, T., Watson, C., Muniz-Medina, V., Christopoulos, A., and Novick, S. (2012) A simple method for quantifying functional selectivity and agonist bias, *ACS Chem Neurosci 3*, 193-203.
[27] Mailman, R. B., Lawler, C.P., Lewis, M.M., Blake, B., Nichols, D.E. (1998) Functional Effects of Novel Dopamine Ligands: Dihydrexidine and Parkinson's Disease as a First Step, In *Dopamine Receptor Subtypes* (Demirdamar, P. J. a. R., Ed.), pp 64-83, IOS Press, Inc, Burke.
[28] Krupnick, J. G., and Benovic, J. L. (1998) The role of receptor kinases and arrestins in G protein-coupled receptor regulation, *Annu Rev Pharmacol Toxicol 38*, 289-319.
[29] Montandon, G., Ren, J., Victoria, N. C., Liu, H., Wickman, K., Greer, J. J., and Horner, R. L. (2016) G-protein-gated Inwardly Rectifying Potassium Channels Modulate Respiratory Depression by Opioids, *Anesthesiology 124*, 641-650.
[30] Hill, R., Disney, A., Conibear, A., Sutcliffe, K., Dewey, W., Husbands, S., Bailey, C., Kelly, E., and Henderson, G. (2018) The novel mu-opioid receptor agonist PZM21 depresses respiration and induces tolerance to antinociception, *Br J Pharmacol 175*, 2653-2661.
[31] Lewis, J. W. (1985) Buprenorphine, *Drug Alcohol Depend 14*, 363-372.
[32] Raffa, R. B., Haidery, M., Huang, H. M., Kalladeen, K., Lockstein, D. E., Ono, H., Shope, M. J., Sowunmi, O. A., Tran, J. K., and Pergolizzi, J. V., Jr. (2014) The clinical analgesic efficacy of buprenorphine, *J Clin Pharm Ther 39*, 577-583.

[33] Warner, N. S., Warner, M. A., Cunningham, J. L., Gazelka, H. M., Hooten, W. M., Kolla, B. P., and Warner, D. O. (2020) A Practical Approach for the Management of the Mixed Opioid Agonist-Antagonist Buprenorphine During Acute Pain and Surgery, *Mayo Clin Proc 95*, 1253-1267.
[34] van Dorp, E., Yassen, A., Sarton, E., Romberg, R., Olofsen, E., Teppema, L., Danhof, M., and Dahan, A. (2006) Naloxone reversal of buprenorphine-induced respiratory depression, *Anesthesiology 105*, 51-57.
[35] Kuo, A., Wyse, B. D., Meutermans, W., and Smith, M. T. (2015) In vivo profiling of seven common opioids for antinociception, constipation and respiratory depression: no two opioids have the same profile, *Br J Pharmacol 172*, 532-548.
[36] Brown, S. M., Campbell, S. D., Crafford, A., Regina, K. J., Holtzman, M. J., and Kharasch, E. D. (2012) P-glycoprotein is a major determinant of norbuprenorphine brain exposure and antinociception, *J Pharmacol Exp Ther 343*, 53-61.
[37] Huang, P., Kehner, G. B., Cowan, A., and Liu-Chen, L. Y. (2001) Comparison of pharmacological activities of buprenorphine and norbuprenorphine: norbuprenorphine is a potent opioid agonist, *J Pharmacol Exp Ther 297*, 688-695.
[38] Sadee, W., Rosenbaum, J. S., and Herz, A. (1982) Buprenorphine: differential interaction with opiate receptor subtypes in vivo, *J Pharmacol Exp Ther 223*, 157-162.
[39] Madia, P. A., Navani, D. M., and Yoburn, B. C. (2012) [(35)S]GTPgammaS binding and opioid tolerance and efficacy in mouse spinal cord, *Pharmacol Biochem Behav 101*, 155-165.
[40] Selley, D. E., Sim, L. J., Xiao, R., Liu, Q., and Childers, S. R. (1997) mu-Opioid receptor-stimulated guanosine-5'-O-(gamma-thio)-triphosphate binding in rat thalamus and cultured cell lines: signal transduction mechanisms underlying agonist efficacy, *Mol Pharmacol 51*, 87-96.
[41] Kenakin, T. P. (2008) Pharmacological onomastics: what's in a name?, *Br J Pharmacol 153*, 432-438.
[42] Ehrlich, A. T., Semache, M., Gross, F., Da Fonte, D. F., Runtz, L., Colley, C., Mezni, A., Le Gouill, C., Lukasheva, V., Hogue, M., Darcq, E., Bouvier, M., and Kieffer, B. L. (2019) Biased Signaling of the Mu Opioid Receptor Revealed in Native Neurons, *iScience 14*, 47-57.
[43] de Waal, P. W., Shi, J., You, E., Wang, X., Melcher, K., Jiang, Y., Xu, H. E., and Dickson, B. M. (2020) Molecular mechanisms of fentanyl mediated beta-arrestin biased signaling, *PLoS Comput Biol 16*, e1007394.
[44] Podlewska, S., Bugno, R., Kudla, L., Bojarski, A. J., and Przewlocki, R. (2020) Molecular Modeling of micro Opioid Receptor Ligands with Various Functional Properties: PZM21, SR-17018, Morphine, and Fentanyl-Simulated Interaction Patterns Confronted with Experimental Data, *Molecules 25*.
[45] Schneider, S., Provasi, D., and Filizola, M. (2016) How Oliceridine (TRV-130) Binds and Stabilizes a mu-Opioid Receptor Conformational State That Selectively Triggers G Protein Signaling Pathways, *Biochemistry 55*, 6456-6466.

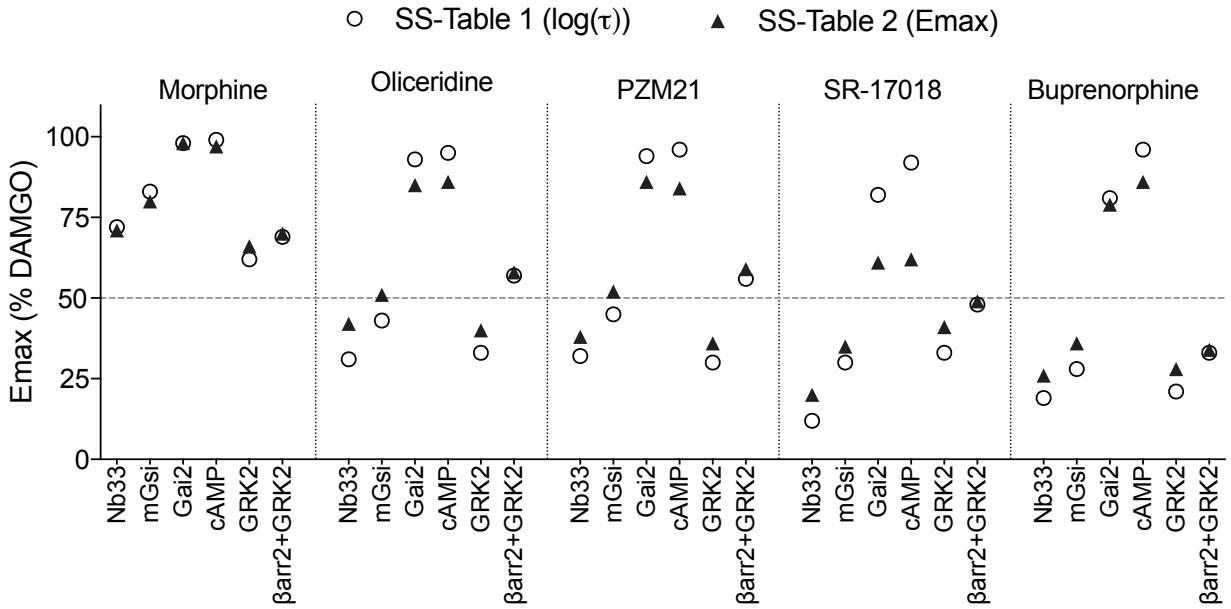


Figure 1. Emax is graphically presented from the values provided in SS-Table 2 in comparison to the Emax calculated from log(τ) (SS-Table 1) using Equation 1.

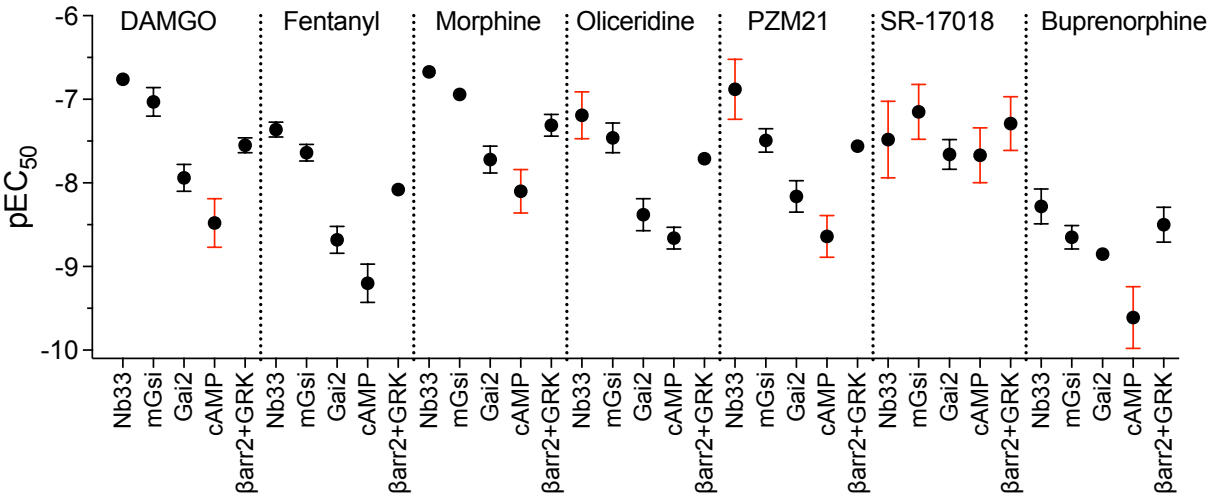


Figure 2. Potency values (pEC_{50}) with standard error of the mean are graphically represented from SS-Table 3. The red error bars indicate S.E.M. values that are one half log order or greater. The SS-manuscript states that means are the average of 3-14 experiments.

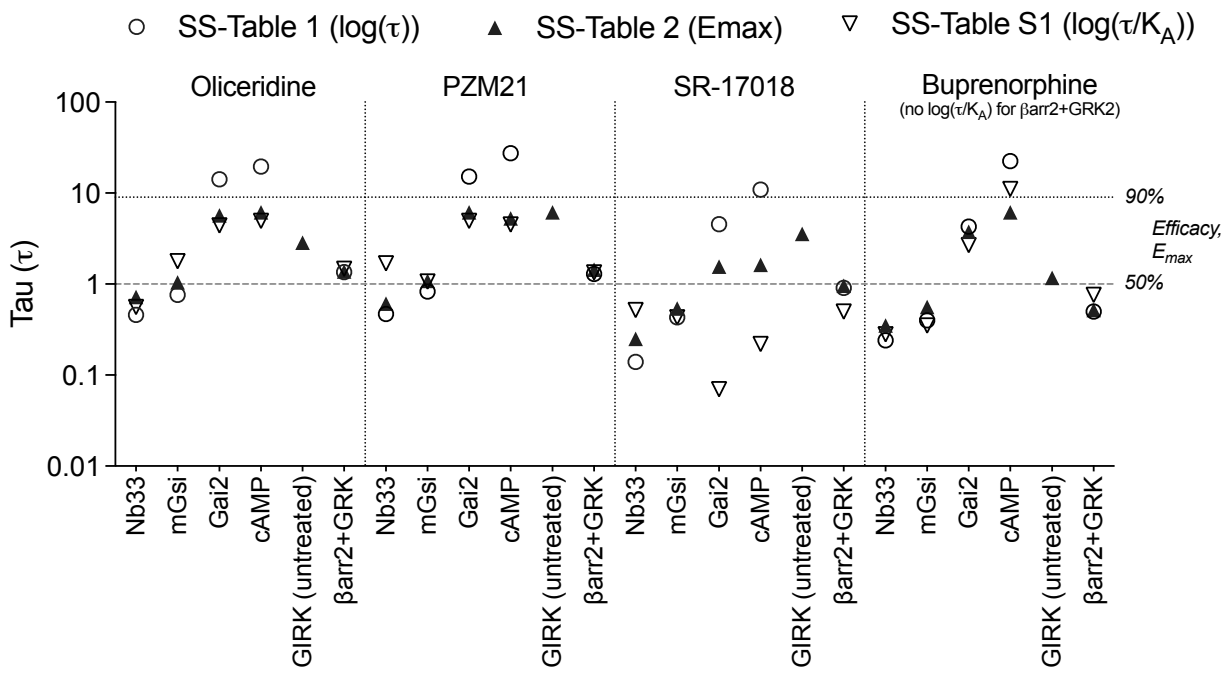


Figure 3. Tau (τ) is graphically presented after derivation from log(τ) (SS-Table 1), Emax (SS-Table 2) or from log(τ/K_A) (SS-Table S1) using Equations 1–6.

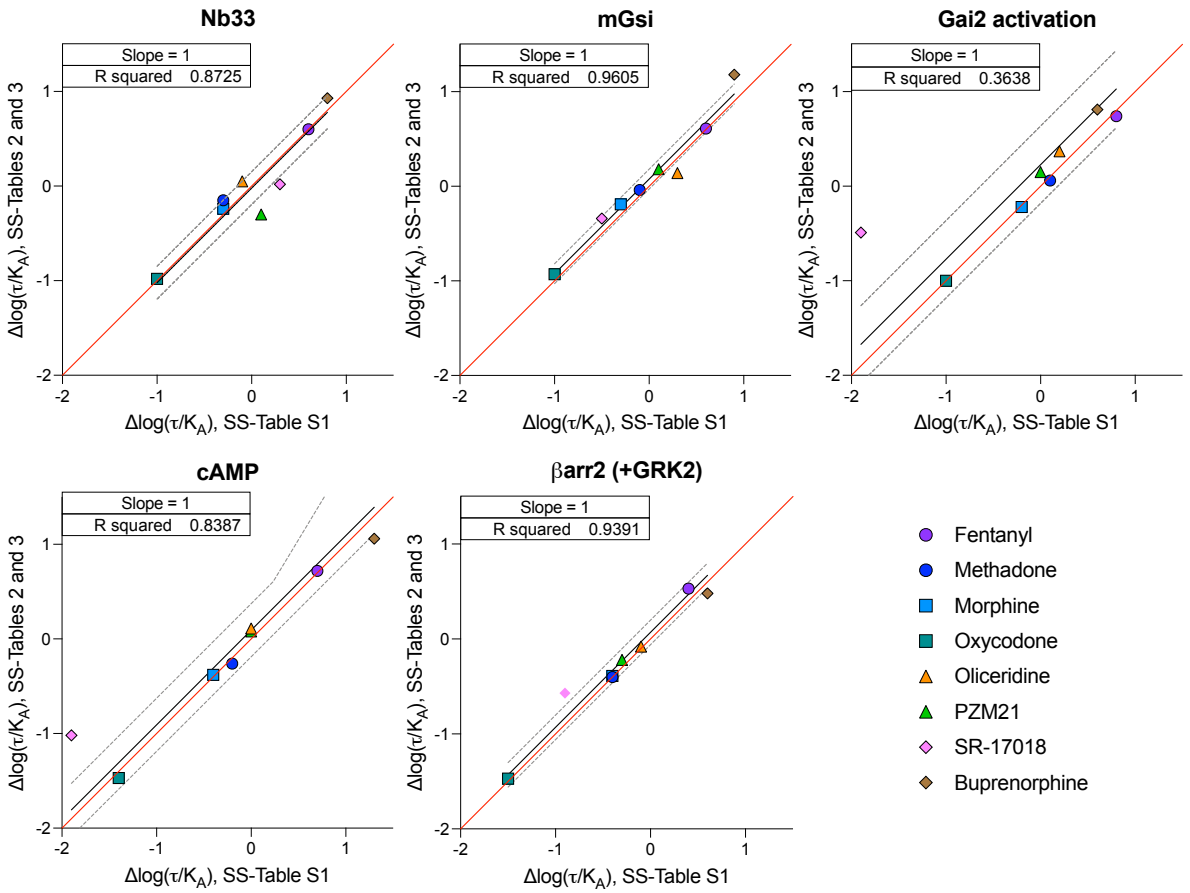
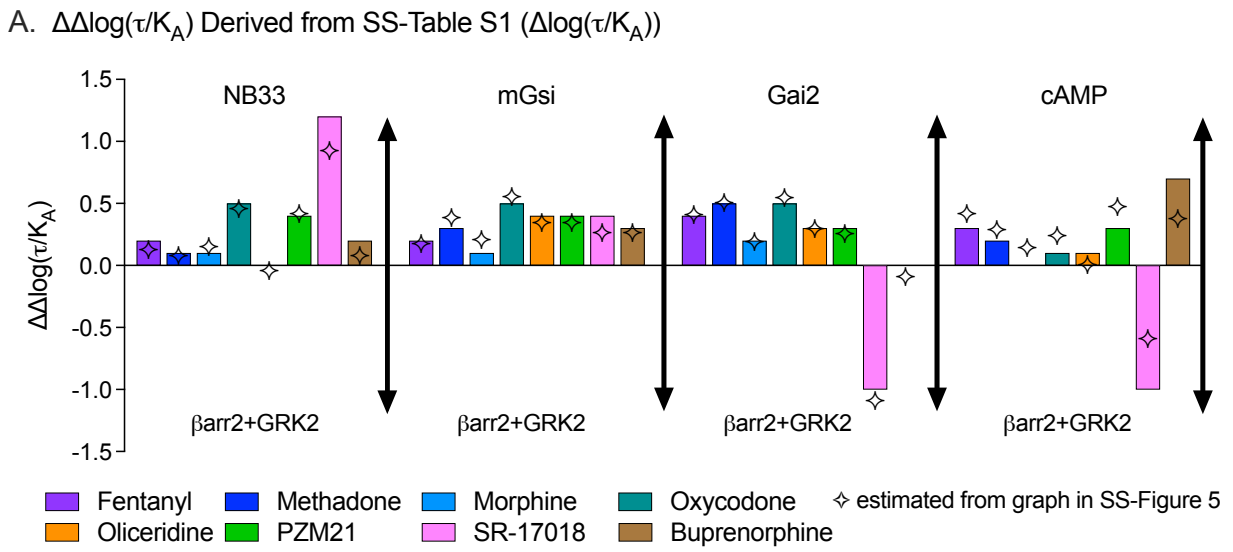


Figure 4. Linear regression of the correlation between the $\Delta\log(\tau/K_A)$ values derived from potency and efficacy values, presented in SS-Tables 2&3, and the $\Delta\log(\tau/K_A)$ values derived from the $\log(\tau/K_A)$ provided in SS-Table S1. The red line indicates a perfect match of the data points (slope=1), and the black line is the linear fit of the data points with 95% CI shown as dashed gray lines.



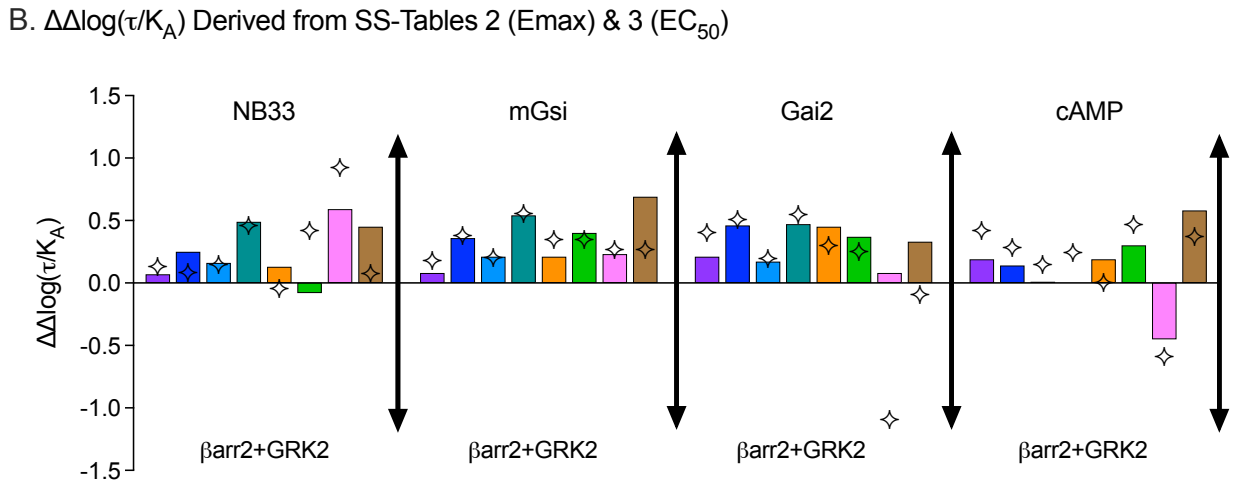


Figure 5. Bias factor analysis presented by calculated $\Delta\Delta\log(\tau/K_A)$ from **A.** SS-Table 1 ($\Delta\log(\tau/K_A)$) or **B.** from $\Delta\log(\tau/K_A)$) values calculated using the potency and efficacy presented in SS-Tables 2&3. The four-point star represents the values published in SS-Figure 5 as determined by graphical overlay trace estimation. The figure legend is the same for A and B.

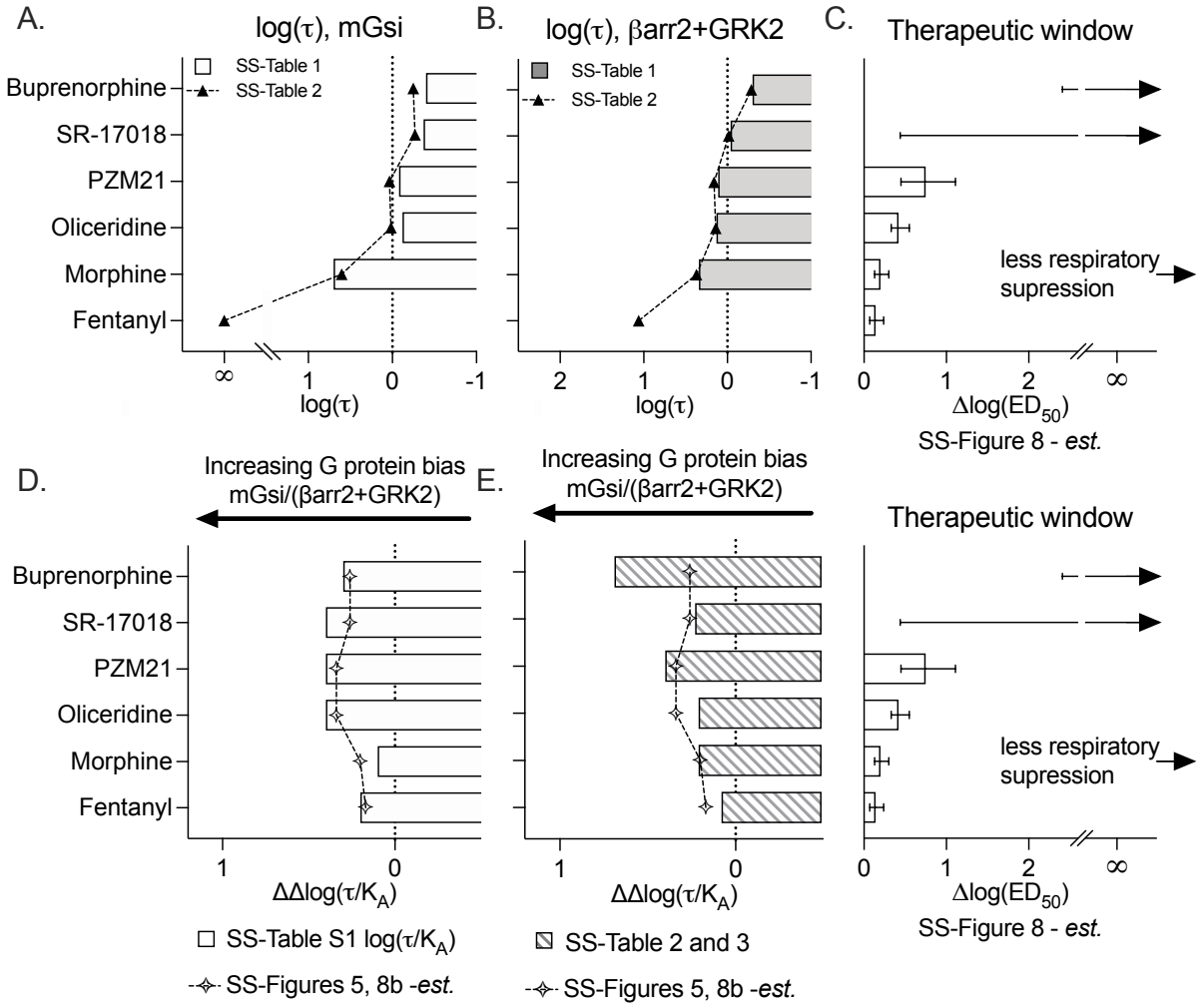


Figure 6. Graphical representation agonist log(τ) and bias factors ($\Delta\Delta\log(\tau/K_A)$ in comparison to therapeutic index as presented in SS-Figure 8. Log(τ) values were plotted from SS-Table 1 or derived from SS-Table 2 and plotted for **A.** mGsi and **B.** βarr2+GRK2 assays and aligned for comparison with the **C.** therapeutic window and error values estimated from SS-Figure 8. Infinity indicates an inability to estimate the mean and error bars are provided as an indication of the lower confidence interval with an arrow pointing towards infinity according to SS-Figure 8 **D**. $\Delta\Delta\log(\tau/K_A)$ values calculated from the $\Delta\log(\tau/K_A)$ values provided in SS-Table S1 are compared to the values presented in SS-Figures 5 and 8B (estimated). **E.** $\Delta\Delta\log(\tau/K_A)$ values derived from EC_{50} (SS-Table 3) and Emax (SS-Table 2) values as described in the text. The therapeutic window plot is shown again for visual alignment.

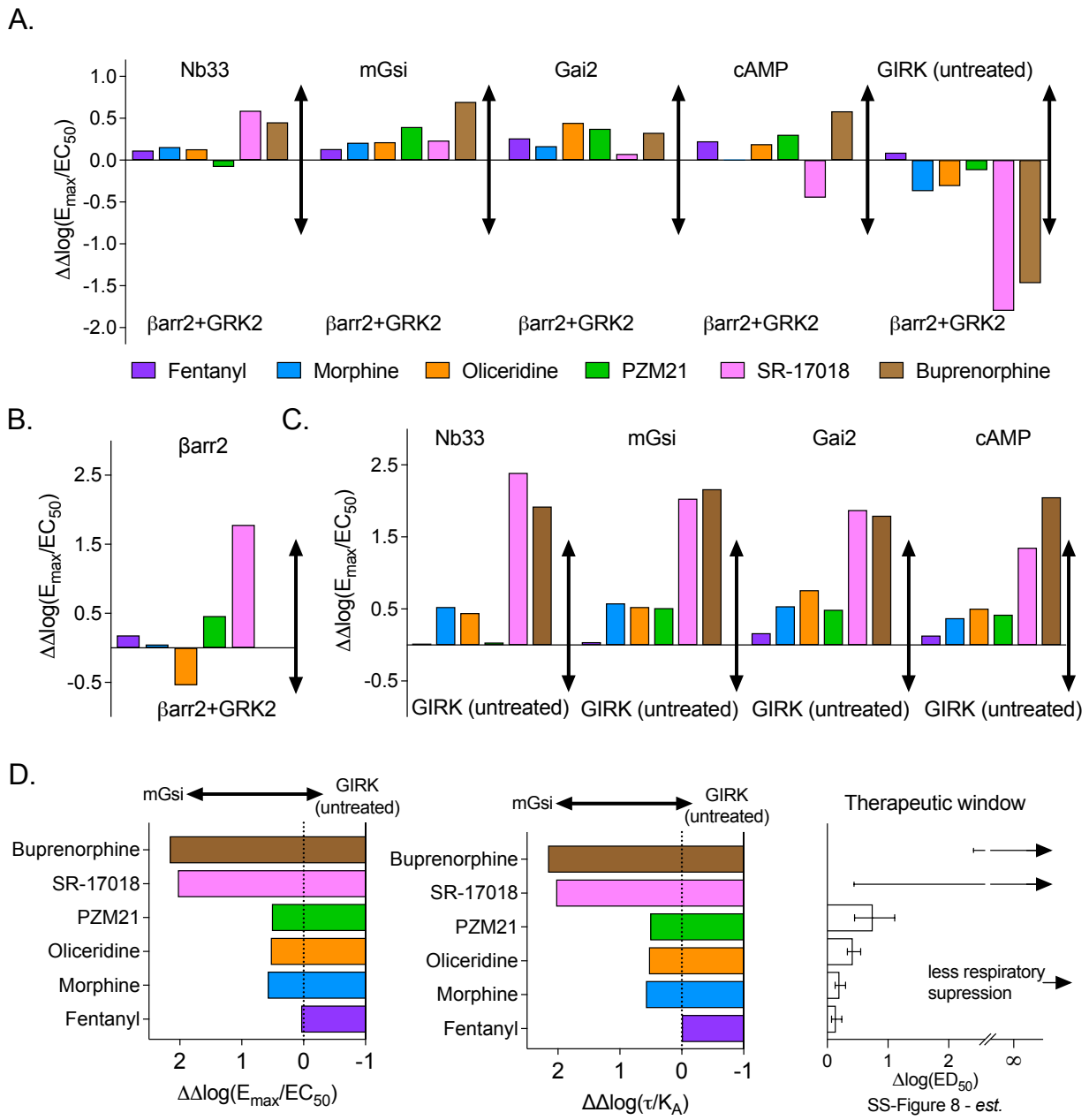


Figure 7. Bias factors determined as $\Delta\Delta\log(Emax/EC_{50})$ from values provided in SS-Tables 2 & 3 comparing responses between **A.** the indicated assays and βarr2 + GRK2 as $\Delta\Delta\log(Emax/EC_{50})_{[G\ protein\ assay\ -\ \beta arr2+GRK2\ assay]}$; **B.** βarr2 vs. βarr2+GRK2 as $\Delta\Delta\log(Emax/EC_{50})_{[\beta arr2\ assay\ -\ \beta arr2+GRK2\ assay]}$; and **C.** between the indicated assays and GIRK (untreated) as $\Delta\Delta\log(Emax/EC_{50})_{[G\ protein\ assay\ -\ GIRK\ untreated\ assay]}$. **D.** Comparison of the as $\Delta\Delta\log(Emax/EC_{50})_{[mGsi\ assay\text{-}GIRK\ assay]}$ and as $\Delta\Delta\log(\tau/K_A)_{[mGsi\ assay\text{-}GIRK\ untreated\ assay]}$ to the therapeutic window from SS-Figure 8 as also shown in Figure 6.

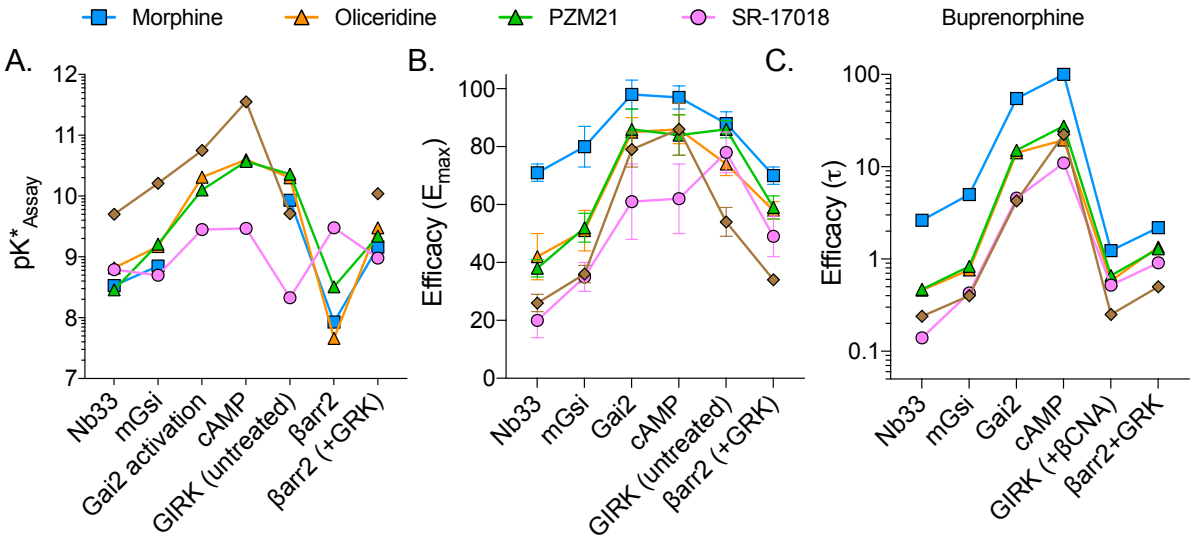


Figure 8. Comparison of intrinsic efficacy of each agonist in the indicated assays as determined by **A.** calculated pK*, (potency values were not included for morphine at Gai and cAMP and buprenorphine at βarr2). **B.** Emax with SEM from SS-Table 2, **C.** τ as shown in Figure 2 as derived from log(τ) (SS-Table 1). Missing values in A for morphine are indicated by interrupted lines. Changes in rank order efficacy are easily visualized by the intersection of lines.