Re-evaluating how low intrinsic efficacy and apparent bias for G protein activation relates to 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 by Gillis et al., 2020 (*1*), 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 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. The designation of *full agonist* is made to compounds that produce the highest observable activation in a system (maximum intrinsic efficacy); agonists producing some fraction of that response are considered *partial agonists*. The maximum response window is determined by the cellular environment, receptor and effector expression levels, and the amplification readout of the system. Biased agonism takes into consideration not only intrinsic efficacy, but also potency (concentration required to reach half maximal efficacy) of an agonist in an assay. Herein, the data published in the aforementioned manuscript was used to rederive the 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 does not support the conclusion that biased agonism, favoring G protein signaling, was not present. Further, 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, introducing G protein signaling bias may be a means to improve opioid analgesia while avoiding certain undesirable side effects.

In the March 31, 2020 issue of Science Signaling, a research article by Gillis et al. (*1*), hereafter referred to as the SS-manuscript) was published that investigated a series of mu opioid ligands and their agonist activity in a compendium of *in vitro* cell-based bioluminescence resonance energy transfer (BRET) studies to investigate the pharmacological basis of functional selectivity (**Table 1**). In particular, the authors focused on responses that involved G protein signaling and βarrestin2 recruitment. Similar BRET-based studies have also been elegantly employed for similar analysis by Ehrlich et al., (*2*) to investigate ligand bias at the mu opioid receptor. The SS-manuscript 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 therefore, concluded that it must be agonist intrinsic efficacy and not biased agonism that 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 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.

In the SS-manuscript, it is proposed that rank order ligand efficacy and *not* biased agonism 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 (*15*). Notably, the authors reanalyzed efficacy (Emax) 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, 16, 17*). *The conclusions of this report* (*1*) *and the value of partial agonist efficacy has been the subject of a number of recent review papers (16-18).*

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 results in 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 $\tau = 1$ (*19*). This can be realized by substitution of $\tau = 1$ into Equation 1, from Black and Leff (*15*):

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*)).

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, 15*)); 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 -0.15 ± 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 (*20*), (*17*). 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.

It is noteworthy that some of the partial agonists produced EC_{50} values that were orders of magnitude apart, across assays, suggesting the possibility of variable active (and/or inactive) state affinity (SS-Tables 3, **Supplemental Figure 1**). It is possible that ligand binding kinetic properties are (mis-)leading to overestimates of potency in certain assays; this would also artificially affect the perception of bias (*21*). 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*).

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 demonstrate 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, $\log(\tau/K_A)$ (*22, 23*). 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 of 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 (*23*). 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 Figure 2,** simulated dose-response curves are included in the appended GraphPad® Prism files).

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 Log(τ/K_A) and K_A known, τ and log(τ) 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 (*17*). This follows from the parameter definition originally described by Black and Leff, (*15*) 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(τ/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 2**. 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 comparison of 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 (and the accompanying GraphPad® Prism files) used for the analyses, herein, are provided for examination in the appendix.

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 $\log(\tau/K_A)$ 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 3,** 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 $\Delta pEC_{50\,Test\text{-}Reference}$. Of note, the column values for Gαi2 recruitment and cAMP inhibition are identical in SS-Table S1. Nonetheless, the $\Delta\log(\tau/K_A)$ for each response was calculated. 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.

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 active state selectivity between one response and another (i.e. functional selectivity, biased agonism) (*23*). 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 (*24*), (*25*).

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 4A** 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 βarrestin2 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 4B**. The

calculations that produced all values determined in **Figure 4** are provided in the supplemental Microsoft® Excel files.

The comparison of $\Delta\Delta\log(\tau/K_A)$ (**Figure 4**) taken from the from 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 (**Figure 4A & 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 an order of magnitude) would not overlap zero as the new mean calculated $\Delta\Delta\log(\tau/K_A)$ changes substantially (by almost two orders 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 by (*23*) was not used, indicating that a direct reassignment to these new calculated mean values would not be possible or useful.

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, $\log(\tau)$, 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 5** presents a representation of the intrinsic efficacy, log(τ), taken from SS-Table S1 or derived from SS-Table 2 (Emax) graphed for mGsi (**Figure 5A**) and βarr2+GRK2 (**Figure 5B**); this is compared with therapeutic window (**Figure 5C**). 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 5B; no value was provided in SS-table 1 and therefore, that bar is represented as approaching infinity. 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 5A-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 5C** was created by essentially 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 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 the mean approaches 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 5C** presents the lower confidence intervals drawn to extend to infinity as depicted in SS-Figure 8.

Figure 5D and 5E are from **Figure 4** and are plotted as presented the SS-manuscript SS-Figure 8, for comparison to the therapeutic window plot (shown again within the figure). **Figure 5D** 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 (no values for buprenorphine were included in SS-Table S1). Here it can be clearly seen that, using the Emax and EC_{50} values to derive $\log(\tau/K_A)$, 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 index.

In this reanalysis of the data presented in the SS-manuscript, efforts were made to rederive τ and K_A from Emax and EC_{50}, presented in the paper, to arrive at the $\Delta\Delta\log(\tau/K_A)$ bias factors. 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 functional selectivity (*22*) in Equation 10:

Equation 10.

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

The results of this relatively simply analysis, can be visualized in **Figure 6,** where a preference for G protein signaling over βarrestin2 (+GRK) is evident across the four platforms. This analysis is in direct agreement with the $\Delta\Delta\log(\tau/K_A)$ values derived from analysis of Emax and EC_{50}, to generate τ and K_A **Figure 5B**.

In the re-calculation of bias from the SS-manuscript it was not possible to interpret error as there was only limited information provided by the authors; for example, number of experiments were provided as a range (3-14 replications) and only 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 **Supplemental Figure 2** of this manuscript, 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 analysis, 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 **(Supplemental 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 (*26*).

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, (*27*). 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 (*28*), as is morphine, even though both perform as partial agonists across many assay systems, as demonstrated in the SS-manuscript. As such, there is not a direct correlation between partial agonism observed in different cell-based assays and the effects of a drug *in vivo*; it may not be fully predictive that a partial agonist may be safer regarding respiratory suppression. Indeed, buprenorphine produces respiratory suppression in humans (*29*), with the additional complication that it is difficult to reverse buprenorphine-induced respiratory suppression with naloxone (*30*) as buprenorphine is a long lasting, high affinity agonist (*27*). 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 (*31*). 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 (*32-34*).

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).

A 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 as 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*). This study 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 (*35*) and in rat thalamus (*36*) in studies assessing ^{35}S-GTPγS binding. Therefore, intrinsic efficacy can be uncoupled from biased agonism and, within that series of agonists, G protein signaling bias was shown to correlate with decreased respiratory suppression, independent of intrinsic efficacy (*8*).

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 (*19, 37*). 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 (*38-40*). 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 CAMYEL	flag-mMOR-NLuc	transient HEK293	BRET
cAMP	cAMP biosensor	YFP-Epac-Rluc	flag-mMOR	transient HEK293	BRET
βarr2+GRK2	βarrestin2 recruitment with overexpressed GRK2	βarr2-YFP (bovine)	flag-MOR-NLuc	transient HEK293	BRET

References.

1. A. Gillis, A. B. Gondin, A. Kliewer, J. Sanchez, H. D. Lim, C. Alamein, P. Manandhar, M. Santiago, S. Fritzwanker, F. Schmiedel, T. A. Katte, T. Reekie, N. L. Grimsey, M. Kassiou, B. Kellam, C. Krasel, M. L. Halls, M. Connor, J. R. Lane, S. Schulz, M. J. Christie, M. Canals, Low intrinsic efficacy for G protein activation can explain the improved side effect profiles of new opioid agonists. *Sci Signal* **13**, (2020); published online EpubMar 31 (10.1126/scisignal.aaz3140).
2. A. T. Ehrlich, M. Semache, F. Gross, D. F. Da Fonte, L. Runtz, C. Colley, A. Mezni, C. Le Gouill, V. Lukasheva, 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 EpubApr 26 (10.1016/j.isci.2019.03.011).
3. 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).
4. L. M. Bohn, R. R. Gainetdinov, F. T. Lin, R. J. Lefkowitz, M. G. Caron, Mu-opioid receptor desensitization by beta-arrestin-2 determines morphine tolerance but not dependence. *Nature* **408**, 720-723 (2000); published online EpubDec 7 (10.1038/35047086).
5. 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).
6. A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hubner, X. P. Huang, M. F. Sassano, P. M. Giguere, S. Lober, D. Da, 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 8 (10.1038/nature19112).
7. N. M. Kennedy, C. L. Schmid, N. C. Ross, K. M. Lovell, Z. Yue, Y. T. Chen, M. D. Cameron, L. M. Bohn, T. D. Bannister, Optimization of a Series of Mu Opioid Receptor (MOR) Agonists with High G Protein Signaling Bias. *J Med Chem* **61**, 8895-8907 (2018); published online EpubOct 11 (10.1021/acs.jmedchem.8b01136).
8. 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**, 1165-1175 e1113 (2017); published online EpubNov 16 (10.1016/j.cell.2017.10.035).
9. 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* **116**, 22353-22358 (2019); published online EpubOct 29 (10.1073/pnas.1908662116).
10. A. Varadi, G. F. Marrone, T. C. Palmer, A. Narayan, M. R. Szabo, V. Le Rouzic, S. G. Grinnell, J. J. Subrath, E. Warner, S. Kalra, A. Hunkele, J. Pagirsky, S. O. Eans, J. M. Medina, J. Xu, Y. X. Pan, A. Borics, G. W. Pasternak, J. P. McLaughlin, S. Majumdar, 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 (2016); published online EpubSep 22 (10.1021/acs.jmedchem.6b00748).

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 (10.1124/jpet.112.201616).
12. D. Lambert, G. Calo, Approval of oliceridine (TRV130) for intravenous use in moderate to severe pain in adults. *Br J Anaesth*, (2020); published online EpubOct 15 (10.1016/j.bja.2020.09.021).
13. S. Ayad, M. A. Demitrack, D. A. Burt, C. Michalsky, L. Wase, M. J. Fossler, A. K. Khanna, 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 (2020); published online EpubAug (10.1007/s40261-020-00936-0).
14. S. Bergese, R. Berkowitz, P. Rider, M. Ladouceur, S. Griffith, A. Segura Vasi, K. Cochrane, L. Wase, M. A. Demitrack, A. S. Habib, Low Incidence of Postoperative Respiratory Depression with Oliceridine Compared to Morphine: A Retrospective Chart Analysis. *Pain Res Manag* **2020**, 7492865 (2020)10.1155/2020/7492865).
15. J. W. Black, P. Leff, Operational models of pharmacological agonism. *Proc R Soc Lond B Biol Sci* **220**, 141-162 (1983); published online EpubDec 22 (10.1098/rspb.1983.0093).
16. A. Gillis, A. Kliewer, E. Kelly, G. Henderson, M. J. Christie, S. Schulz, M. Canals, Critical Assessment of G Protein-Biased Agonism at the mu-Opioid Receptor. *Trends Pharmacol Sci*, (2020); published online EpubOct 20 (10.1016/j.tips.2020.09.009).
17. A. Gillis, V. Sreenivasan, M. J. Christie, Intrinsic Efficacy of Opioid Ligands and Its Importance for Apparent Bias, Operational Analysis, and Therapeutic Window. *Mol Pharmacol* **98**, 410-424 (2020); published online EpubOct (10.1124/mol.119.119214).
18. A. Faouzi, B. R. Varga, S. Majumdar, Biased Opioid Ligands. *Molecules* **25**, (2020); published online EpubSep 16 (10.3390/molecules25184257).
19. T. Kenakin, A. Christopoulos, Signalling bias in new drug discovery: detection, quantification and therapeutic impact. *Nat Rev Drug Discov* **12**, 205-216 (2013); published online EpubMar (10.1038/nrd3954).
20. E. Kelly, Efficacy and ligand bias at the mu-opioid receptor. *Br J Pharmacol* **169**, 1430-1446 (2013); published online EpubAug (10.1111/bph.12222).
21. 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).
22. M. T. Griffin, K. W. Figueroa, S. Liller, F. J. Ehlert, 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 (2007); published online EpubJun (10.1124/jpet.107.120857).
23. 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).
24. K. A. Berg, S. Maayani, J. Goldfarb, C. Scaramellini, P. Leff, W. P. Clarke, 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 (1998); published online EpubJul (

25. R. B. Mailman, Lawler, C.P., Lewis, M.M., Blake, B., Nichols, D.E., in *Dopamine Receptor Subtypes,* P. J. a. R. Demirdamar, Ed. (IOS Press, Inc, Burke, 1998), pp. 64-83.
26. 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 EpubJul (10.1111/bph.14224).
27. J. W. Lewis, Buprenorphine. *Drug Alcohol Depend* **14**, 363-372 (1985); published online EpubFeb (10.1016/0376-8716(85)90067-5).
28. 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).
29. N. S. Warner, M. A. Warner, J. L. Cunningham, H. M. Gazelka, W. M. Hooten, B. P. Kolla, D. O. Warner, A Practical Approach for the Management of the Mixed Opioid Agonist-Antagonist Buprenorphine During Acute Pain and Surgery. *Mayo Clin Proc* **95**, 1253-1267 (2020); published online EpubJun (10.1016/j.mayocp.2019.10.007).
30. E. van Dorp, A. Yassen, E. Sarton, R. Romberg, E. Olofsen, L. Teppema, M. Danhof, A. Dahan, Naloxone reversal of buprenorphine-induced respiratory depression. *Anesthesiology* **105**, 51-57 (2006); published online EpubJul (10.1097/00000542-200607000-00012).
31. A. Kuo, B. D. Wyse, W. Meutermans, M. T. Smith, 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 (2015); published online EpubJan (10.1111/bph.12696).
32. S. M. Brown, S. D. Campbell, A. Crafford, K. J. Regina, M. J. Holtzman, E. D. Kharasch, P-glycoprotein is a major determinant of norbuprenorphine brain exposure and antinociception. *J Pharmacol Exp Ther* **343**, 53-61 (2012); published online EpubOct (10.1124/jpet.112.193433).
33. 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 (
34. W. Sadee, J. S. Rosenbaum, A. Herz, Buprenorphine: differential interaction with opiate receptor subtypes in vivo. *J Pharmacol Exp Ther* **223**, 157-162 (1982); published online EpubOct (
35. P. A. Madia, D. M. Navani, B. C. Yoburn, [(35)S]GTPgammaS binding and opioid tolerance and efficacy in mouse spinal cord. *Pharmacol Biochem Behav* **101**, 155-165 (2012); published online EpubMar (10.1016/j.pbb.2011.11.001).
36. D. E. Selley, L. J. Sim, R. Xiao, Q. Liu, S. R. Childers, 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 (1997); published online EpubJan (10.1124/mol.51.1.87).
37. T. P. Kenakin, Pharmacological onomastics: what's in a name? *Br J Pharmacol* **153**, 432-438 (2008); published online EpubFeb (10.1038/sj.bjp.0707407).
38. P. W. de Waal, J. Shi, E. You, X. Wang, K. Melcher, Y. Jiang, H. E. Xu, B. M. Dickson, Molecular mechanisms of fentanyl mediated beta-arrestin biased signaling. *PLoS Comput Biol* **16**, e1007394 (2020); published online EpubApr (10.1371/journal.pcbi.1007394).
39. S. Podlewska, R. Bugno, L. Kudla, A. J. Bojarski, R. Przewlocki, 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**, (2020); published online EpubOct 12 (10.3390/molecules25204636).
40. S. Schneider, D. Provasi, M. Filizola, How Oliceridine (TRV-130) Binds and Stabilizes a mu-Opioid Receptor Conformational State That Selectively Triggers G Protein Signaling Pathways. *Biochemistry* **55**, 6456-6466 (2016); published online EpubNov 22 (10.1021/acs.biochem.6b00948).

Figure 1

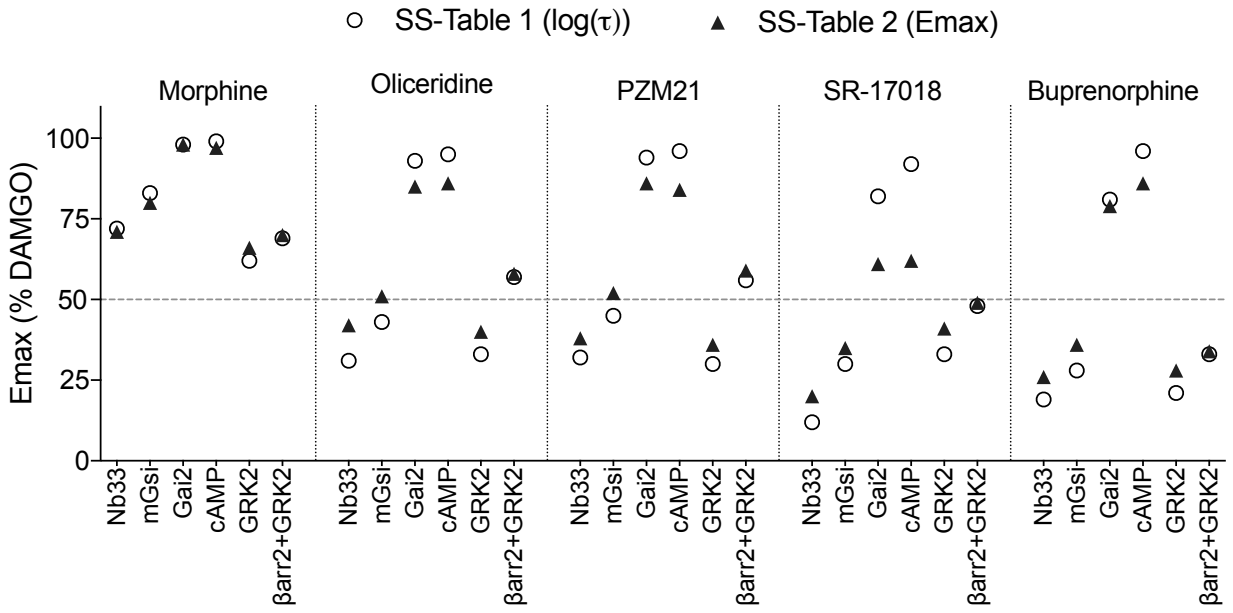


Figure 1. Emax is graphically presented from the values provided in SS-Table 2 in comparison to the Emax calculated from log(τ) (SS-Table 1) using Equation 1.

Figure 2

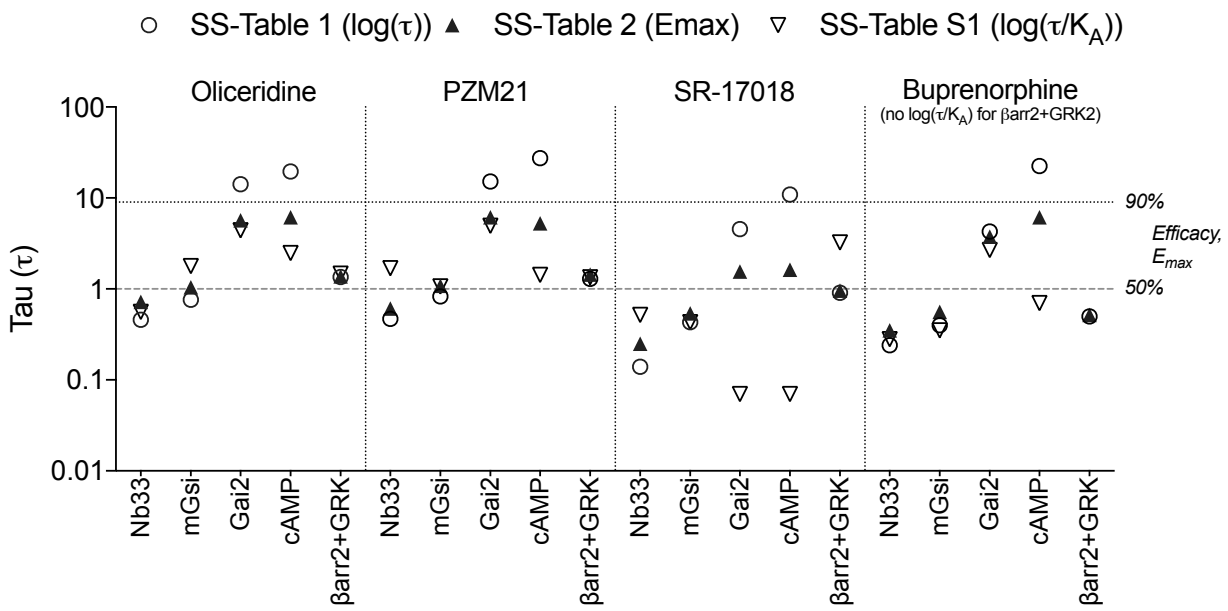


Figure 2. Tau (τ) is graphically presented after derivation from log(τ) (SS-Table 1), Emax (SS-Table 2) or from log(τ/K_A) (SS-Table S1) using Equations 1–6.

Figure 3.

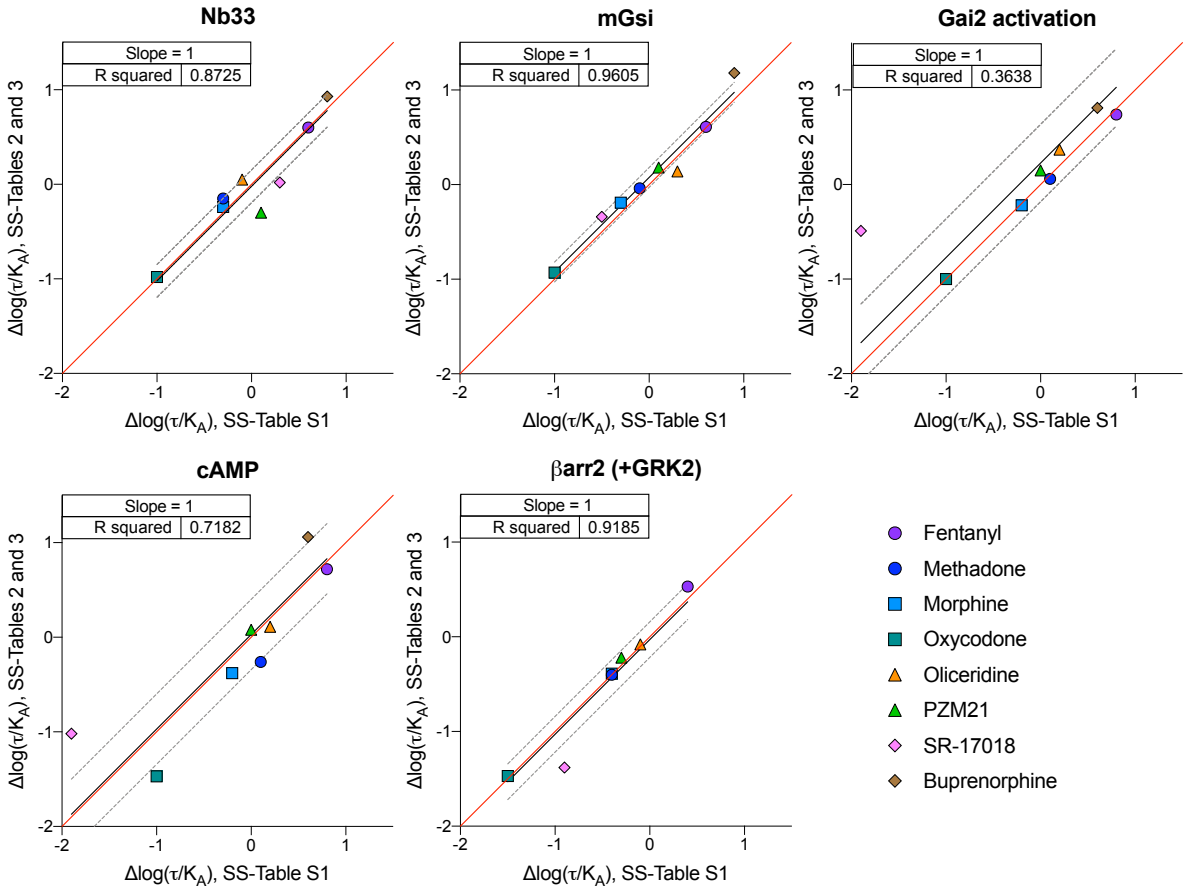


Figure 3. Linear regression of the correlation between the $\Delta\log(\tau/K_A)$ values derived from potency and efficacy values, presented in SS-Tables 2&3, and the $\Delta\log(\tau/K_A)$ values derived from the $\log(\tau/K_A)$ provided in SS-Table S1. The red line indicates a perfect match of the 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 4.

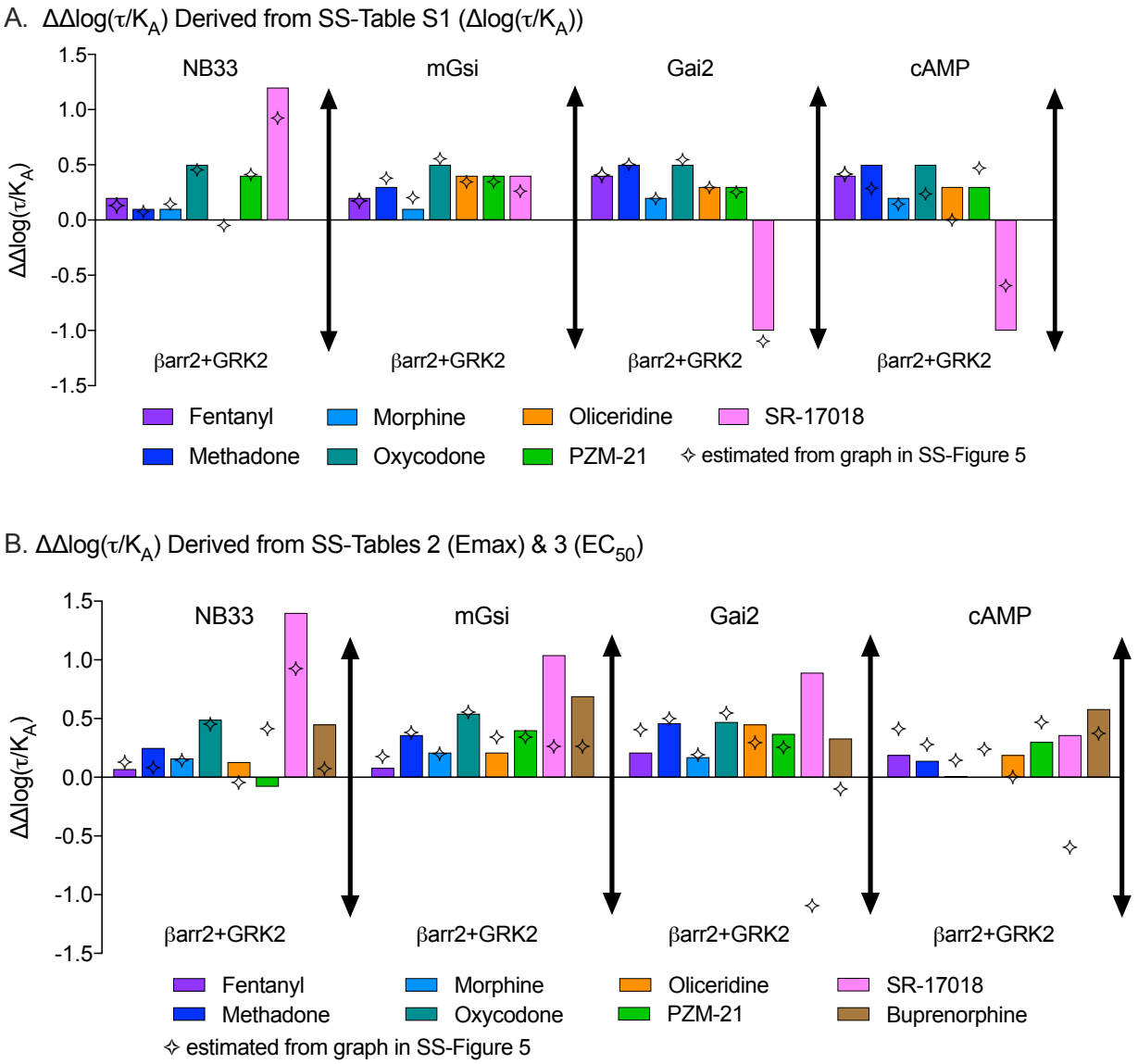


Figure 4. Bias factor analysis presented by calculated $\Delta\Delta\log(\tau/K_A)$ from **A.** SS-Table 1 ($\Delta\log(\tau/K_A)$) or **B.** from $\Delta\log(\tau/K_A)$) values calculated using the potency and efficacy presented in SS-Tables 2&3. The four-point star represents the values published in SS-Figure 5 as determined by graphical overlay estimation.

Figure 5.

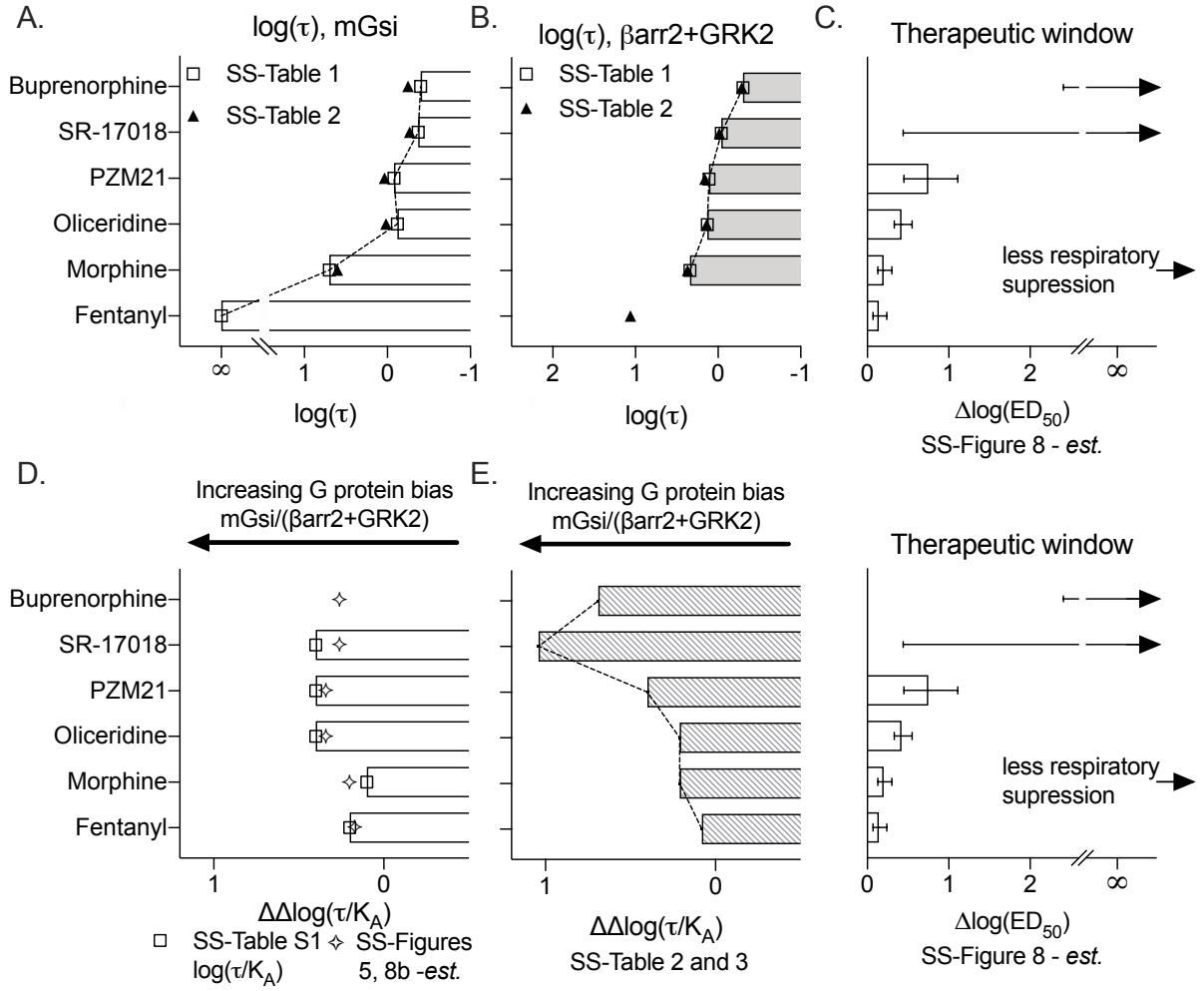


Figure 5. Graphical representation agonist log(τ) and bias factors (ΔΔlog(τ/K_A) in comparison to therapeutic index as presented in SS-Figure 8. Log(τ) values were plotted from SS-Table 1 or derived from SS-Table 2 and plotted for **A.** mGsi and **B.** βarr2+GRK2 assays and aligned for comparison with the **C.** therapeutic window and error values estimated from SS-Figure 8. Infinity indicates and 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.** ΔΔlog(τ/K_A) values calculated from the Δlog(τ/K_A) values provided in SS-Table S1 are compared to the values presented in SS-Figures 5 and 8B (estimated). **E.** ΔΔlog(τ/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 6.

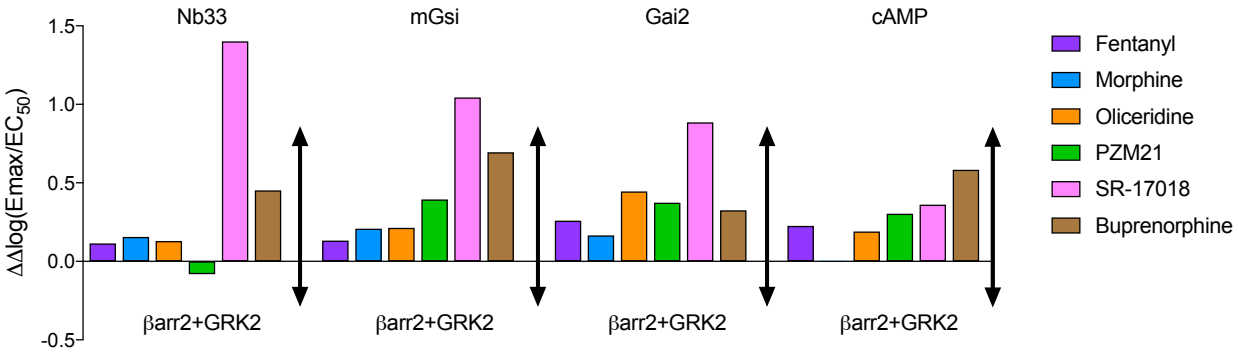
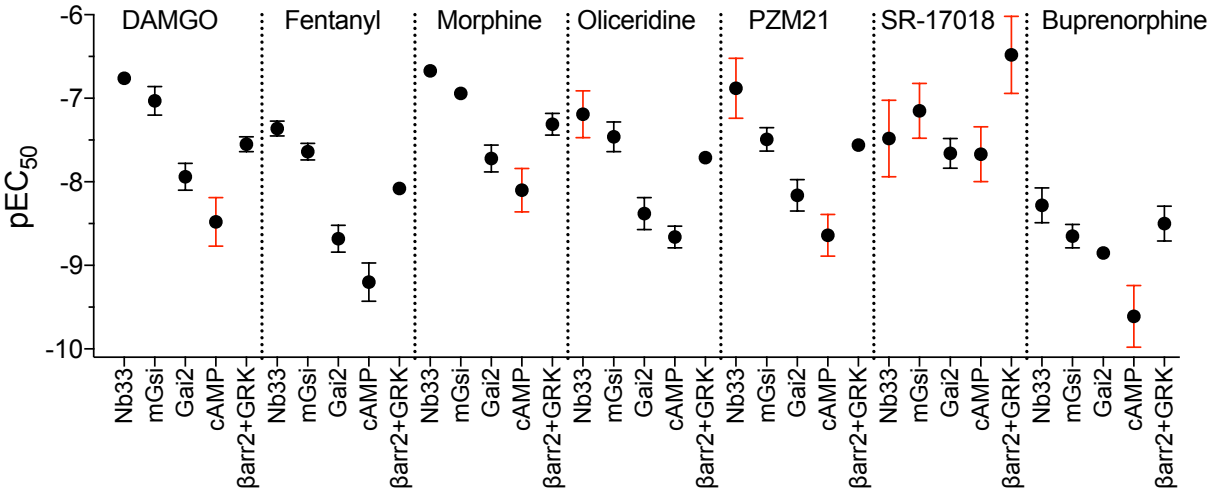


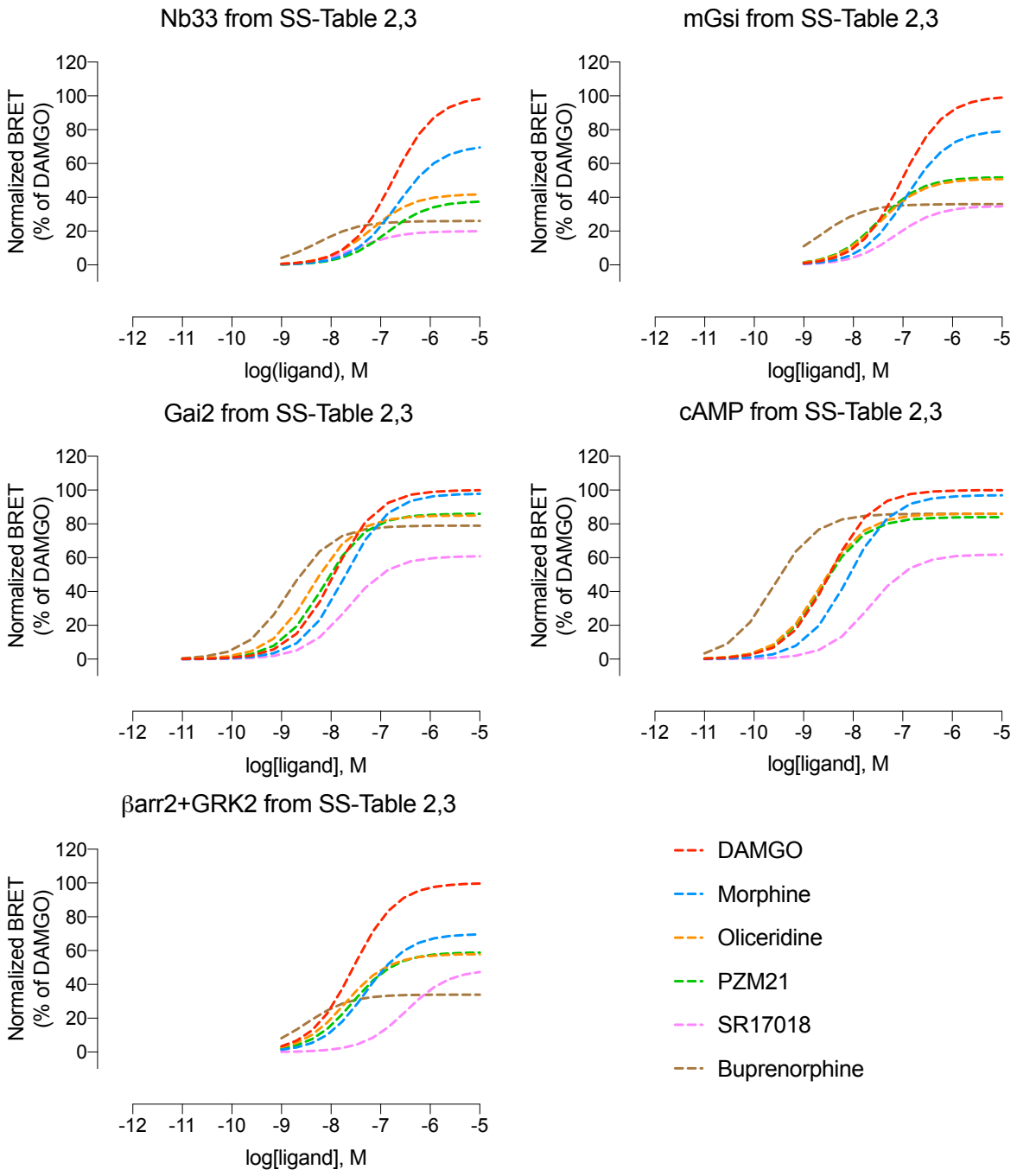
Figure 6. Bias factors, $\Delta\Delta\log(Emax/EC_{50})_{[G\ protein\ assay\ -\ \beta arr2+GRK2\ assay]}$, derived from the Emax and EC_{50} values provided in SS-Tables 2 & 3.

Supplemental Figure 1



Supplemental Figure 1. 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.

Supplemental Figure 2



Supplemental Figure 2. Simulated concentration response curves generated from SS-Table 2 and SS-Table 3 parameters using 3 parameter nonlinear regression, GraphPad Prism 8 (files are appended).