High resolution deep mutational scanning of the melanocortin-4 receptor enables target characterization for drug discovery

Authors: Conor J. Howard[1,#], Nathan S. Abell[1,#], Beatriz A. Osuna[1,4], Eric M. Jones[1], Leon Y. Chan[1], Henry Chan[1], Dean R. Artis[1,5], Jonathan B. Asfaha[1], Joshua S. Bloom[2,3], Aaron R. Cooper[1,6], Andrew Liao[1], Eden Mahdavi[1], Nabil Mohammed[1], Alan L. Su[1,7], Giselle A. Uribe[1], Sriram Kosuri[1], Diane E. Dickel[1,*], Nathan B. Lubock[1,*]

Affiliations:

[1] Octant, Inc., Emeryville, CA, USA
[2] Department of Human Genetics and Department of Computational Medicine, University of California, Los Angeles, Los Angeles CA, USA
[3] Howard Hughes Medical Institute, Chevy Chase, MD, USA
[4] Current affiliation: Tacit Therapeutics, Inc., South San Francisco, CA, USA
[5] Current affiliation: Annexon Biosciences, Brisbane, CA, USA
[6] Current affiliation: Arsenal Biosciences, South San Francisco, CA, USA
[7] Current affiliation: Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305

[#] These authors contributed equally to this work.
[*] These authors contributed equally to this work and to whom correspondence should be addressed: nate@octant.bio (NBL), diane@octant.bio (DED)

Abstract

Deep Mutational Scanning (DMS) is an emerging method to systematically test the functional consequences of thousands of sequence changes to a protein target in a single experiment. Because of its utility in interpreting both human variant effects and protein structure-function relationships, it holds substantial promise to improve drug discovery and clinical development. However, applications in this domain require improved experimental and analytical methods. To address this need, we report novel DMS methods to precisely and quantitatively interrogate disease-relevant mechanisms, protein-ligand interactions, and assess predicted response to drug treatment. Using these methods, we performed a DMS of the melanocortin-4 receptor (MC4R), a G protein-coupled receptor (GPCR) implicated in obesity and an active target of drug development efforts. We assessed the effects of >6,600 single amino acid substitutions on MC4R's function across 18 distinct experimental conditions, resulting in >20 million unique measurements. From this, we identified variants that have unique effects on MC4R-mediated $G\alpha_s$- and $G\alpha_q$-signaling pathways, which could be used to design drugs that selectively bias MC4R's activity. We also identified pathogenic variants that are likely amenable to a corrector therapy. Finally, we functionally characterized structural relationships that distinguish the binding of peptide versus small molecule ligands, which could guide compound optimization. Collectively, these results demonstrate that DMS is a powerful method to empower drug discovery and development.

Introduction

Deep Mutational Scanning (DMS) employs cutting-edge synthetic biology or genome editing methods, DNA synthesis, and sequencing to systematically assess the effect of every possible single amino acid substitution on the function of a protein target (Starita et al. 2017; Araya and Fowler 2011; Fowler and Fields 2014). Researchers have leveraged DMS to gauge a variety of protein functions or their cellular consequences, including viability (Findlay et al. 2014), protein abundance (Matreyek et al. 2018; Faure et al. 2022), transcriptional signaling (Jones et al. 2020), and inter- and intramolecular interactions (Braberg et al. 2022; Faure et al. 2022). DMS assays are increasingly used for human variant interpretation (Weile and Roth 2018) and to elucidate the relationship between protein structure and function, including in the evaluation (Brandes et al. 2023) and finetuning of protein language models (Lafita et al. 2024).

While DMS has significantly advanced our understanding of protein function, its potential in drug discovery and development has yet to be fully realized. Applications in this realm require methods that are more disease-relevant, sensitive, and quantitative to measure the subtle effects of both sequence variants and experimental conditions (e.g., drug treatments). For example, DMS assays often measure effects like viability, which are several biological layers removed from the specific molecular mechanisms modulated by drugs. Additionally, the current signal-to-noise ratios of many assays and challenges around uncertainty quantification make drawing quantitative conclusions from DMS data difficult. Moving from categorical (e.g., benign vs. pathogenic) classifications toward precise quantitative measurements of human variants could improve predictions of safety and efficacy at early stages of drug development programs (Plenge, Scolnick, and Altshuler 2013). It would also expand the use of matching patients' medications to their specific genetic variants (i.e., theratyping), which could lead to better patient outcomes (McDonald et al., 2024). Finally, DMS data contain the functional consequences of thousands of different biochemical perturbations. Sufficiently sensitive assays against functional

readouts and improved analysis methods would readily complement structure-based approaches by elucidating the functional consequences of ligand binding. Building on this, it should be possible to identify novel protein-ligand interactions that could be reverse-engineered to increase compound potency, further expanding the potential impact of DMS on drug discovery.

In a previous work, we described DMS methods to measure the effects of thousands of single amino acid variants on the function of the beta-2 adrenergic receptor (β2AR) (Jones et al. 2020), a member of the G-protein-coupled receptor (GPCR) class that is the most commonly targeted protein family in drug development (Sriram and Insel 2018). Here, we build upon this previous work to demonstrate more sensitive and robust DMS methods for drug discovery and development, focusing on the melanocortin-4 receptor (MC4R). MC4R is a GPCR, and human variants that result in partial or complete loss of its function cause the most common form of inherited obesity [OMIM #618406] (Farooqi et al. 2003; Vaisse et al. 2000; Hinney et al. 2003). Variants that increase MC4R activity are protective against obesity (Lotta et al. 2019; Paisdzior et al. 2020), and numerous small molecule and peptide agonists of MC4R have been tested as potential therapeutics (Sweeney et al. 2023; Greenfield et al. 2009; K. Y. Chen et al. 2015; Kievit et al. 2013; Collet et al. 2017; Clément et al. 2020; Huang and Tao 2014; Hinney, Körner, and Fischer-Posovszky 2022).

In this study, we developed substantially improved experimental and analytical methods for DMS that are capable of detecting subtle quantitative effects of variants and differences between experimental conditions with a high degree of statistical rigor. We then tested the effects of nearly all possible single amino acid variants of MC4R (6,633 of 6,640) on two distinct GPCR signaling functions under a variety of treatment conditions. From this, we generated a high-resolution map of how MC4R's structure relates to function, and we accurately classified

the quantitative effects of human variants. Additionally, we identified amino acid changes that differentially impact (i.e., bias) MC4R's different GPCR signaling functions, pinpointed human variants that are amenable to a specific class of therapy, and elucidated the functional impact of protein-ligand interactions between MC4R and both peptide and small molecule ligands. This demonstrates the utility of DMS for various drug discovery and development applications, and the methods described herein should be broadly applicable to GPCRs and other drug target classes that function in transcriptional signaling pathways.

Results

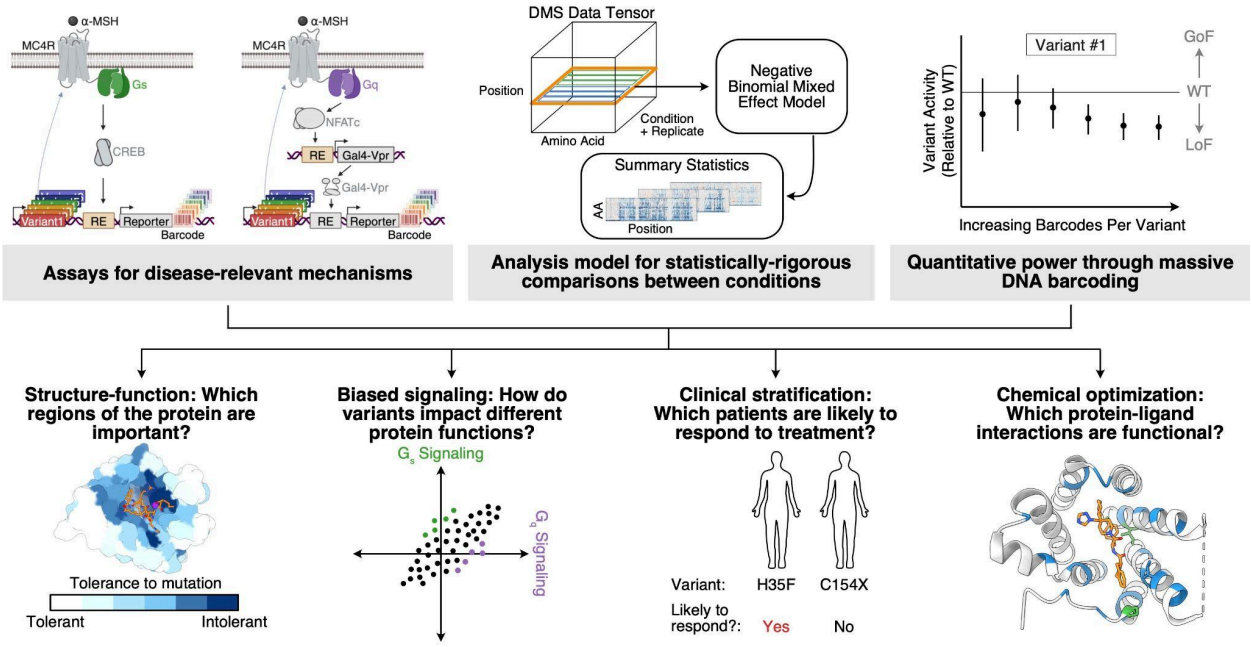


Figure 1. Highly quantitative Deep Mutational Scanning (DMS) methods for drug discovery applications. Top: DMS experimental and analysis methods improvements reported in this paper. **Bottom:** Questions commonly encountered in drug discovery and development that can be addressed by improved DMS methodology.

Development of highly quantitative deep mutational scanning methods

Assays for disease-relevant mechanisms

A critical tool in drug discovery programs is a very sensitive set of assays that directly interrogate the function(s) of a protein target against which to test potential therapies (Hughes et al. 2011). This is essential for ensuring that compounds are modulating a specific mechanism underlying disease and not having undesired “off-target” effects. Stimulation of MC4R with its agonist, alpha melanocyte-stimulating hormone (α-MSH), results in signaling through multiple canonical GPCR pathways, including $G\alpha_s$-coupled cyclic adenosine monophosphate signaling (hereafter referred to as Gs) and $G\alpha_q$-coupled calcium signaling (hereafter referred to as Gq) (Tao 2010). Therefore, we first developed multiplexed reporter assays for these two critical MC4R G-protein signaling functions (**Fig. 1**, **Supplementary Fig. 1**), building off of our earlier

work performing a deep mutational scan of β2AR (Jones et al. 2020). Both reporters were designed for use in human HEK293T cells and to be compatible with our previously described DMS library construction methods (Jones et al. 2020). Briefly, these methods harness high-throughput DNA synthesis to construct every possible single amino acid variant, and each variant is then linked to a transcriptional reporter containing an oligonucleotide sequence barcode unique to that variant. Reporter constructs are then integrated into cells using a site-specific recombination-based landing pad system and drug selection to ensure that each cell contains a single variant-barcode combination. Activation of the receptor turns on a response element for the signaling pathway, leading to the expression of the barcoded reporter, which is then quantified using RNA sequencing.

The MC4R Gs assay was adapted and further optimized from the cAMP response element-based reporter we previously used for β2AR (**Supplementary Fig. 1A,C**). For Gq signaling, an analogous reporter using an NFAT response element (Boss, Talpade, and Murphy 1996) alone to activate reporter gene expression was not suitable due to weak signal-to-noise (**Supplementary Fig. 1D**). To solve this problem, we incorporated a “relay” system to amplify the reporter signal using a synthetic transcription factor composed of Gal4 fused to the VP64-p65-Rta (VPR) transcriptional activator (Chavez et al. 2015) (**Fig. 1**, **Supplementary Fig. 1B**, see **Methods** for more details). The resulting Gal4-VPR transcription factor was placed under the control of the NFAT response element, and the reporter gene was placed under the control of a UAS element that responds to Gal4 binding. This assay design resulted in robust reporter expression upon stimulation of MC4R (**Supplementary Fig. 1C**). Together, these assays provide sensitive functional readouts for two important MC4R signaling activities implicated in obesity and other phenotypes (Fatima et al. 2022; Sweeney et al. 2023), and we used them to build DMS libraries to assess the effects of all possible single amino acid substitutions in MC4R.

Analysis model for statistically robust comparisons

To explore many hypotheses that arise in drug discovery applications, it is valuable to assay a DMS library using experimental replication and under a variety of conditions, such as different drugs and/or pathways of interest. However, most popular methods for DMS analysis do not leverage DNA barcodes or other experimental replicate information, nor do they support hypothesis testing between conditions (Rubin et al. 2017; Faure et al. 2020). Consequently, we developed an alternative modeling framework to enable this (**Fig. 1, Supplementary Fig. 2A**, see **Methods: Negative Binomial Regression Analysis Pipeline** for additional details). Briefly, borrowing from approaches for inferring differential expression from RNA-seq data (McCarthy, Chen, and Smyth 2012; Love, Huber, and Anders 2014; Ahlmann-Eltze and Huber 2021), we applied a mixed effect negative binomial generalized linear model (GLM) to raw barcode counts directly. The model contains a random effect across barcodes to share barcode information between replicates and conditions, and incorporates sample-specific offsets to account for technical covariates like sequencing depth, as is common for RNA-seq (Robinson and Oshlack 2010). For each variant, we estimate the mean shift in barcode count and associated standard error for each treatment condition, relative to wild-type. Using per-condition summary statistics, we either directly test whether each variant barcode mean is significantly different from wild-type (zero) in each treatment, or we can define more complex linear contrasts on variant effects across multiple treatments.

Increasing power through barcoding

As described above, our assay design and analysis framework harness DNA barcodes that are uniquely associated with a particular variant and provide multiple independent measures of a variant's effect. In our previous DMS of β_2AR, the median number of barcodes independently linked to each variant was ~10 (Jones et al. 2020), and we reasoned that increasing this number

would increase the power to detect functional effects. To this end, we optimized and scaled our library cloning and cellular integration protocols to target ~30 barcodes per variant in building DMS libraries for MC4R. As expected, this increased the power to detect variant effects. For example, the separation between the activity of alleles that are clearly deleterious (i.e., a stop codon at any position in the protein) and all other alleles (i.e., wild-type or missense variants) was drastically increased in the MC4R Gs assay relative to the same assay for β_2AR (**Supplementary Fig. 2B**). To further test the effect of the number of barcodes for a given variant, we computationally down-sampled the barcodes for representative positions of MC4R and ran the resulting data through our analysis pipeline. As expected, by increasing the number of barcodes per variant, the magnitude of the standard error of the estimated variant effect decreases in a manner consistent with increased sample size (**Supplementary Fig. 2C**). This confirms that increasing the number of barcodes per variant enables the quantification of subtle differences within a standard hypothesis testing framework, and provides an experimental parameter that one can vary to improve the power to detect the effects of sequence variants of particular interest.

Comprehensive deep mutational scanning of MC4R

With these methods in hand, we carried out a comprehensive assessment of the effects of all single amino acid substitutions (including nonsense variants) on MC4R's Gs and Gq signaling activities under a variety of experimental conditions (**Fig. 2a**, **Supplementary Table 1, Supplementary Figs. 3-4**). We selected experimental conditions that would inform aspects of drug discovery and development programs, such as elucidating protein structure-function relationships, identifying regions of the protein that bias activity towards or away from a specific function, classifying the effects of human variants in the presence and absence of potential therapies, and uncovering functional differences in protein-ligand interactions. In total, we tested 18 unique conditions, each performed in quadruplicate, including: basal activity (i.e., no

stimulation) of MC4R, stimulation of MC4R with a range of doses of the native peptide agonist alpha-melanocyte-stimulating hormone (α-MSH), stimulation with a range of doses of a small molecule agonist (THIQ) (Sebhat et al. 2002), treatment with a small molecule corrector (Ipsen 17) (Wang et al. 2014; Poitout et al. 2007), and library composition normalization controls (forskolin, see **Methods**). Our resulting DMS assays had extraordinary variant coverage, with 99.9% (6,633/6,640) of all possible single amino acid substitutions present in all experimental conditions. Each variant was represented by an average of 56 and 28 barcodes for the Gs and Gq signaling pathways, respectively. Between both assays, this translates to more than 557,000 uniquely engineered human cells, each containing a distinctive variant-reporter-barcode combination. When factoring in the number of experimental conditions (18 unique), replicates (four per condition), amino acid variants tested (99.9% of 6,640 possible), and the mean barcodes per variant (56 and 28 for Gs and Gq, respectively), this equates to >21,500,000 measurements across all datasets.

Multiple lines of evidence support the high quality and utility of these data for classification of variant effects (**Fig. 2A-F, Supplementary Figs. 3-5**). Focusing on one dataset as a representative example (Gs signaling using a low dose of α-MSH stimulation), variants that introduce stop codons or fall within transmembrane domains and buried surfaces disproportionately lead to significant loss of MC4R function (**Fig. 2A-C**). The results also correlate well with expectations from human genetics data and variant effect prediction algorithms (**Fig. 2C-D**, **Supplementary Fig. 5**). For example, the majority (63.3%, 31/49) of human MC4R variants classified as pathogenic or likely pathogenic in ClinVar (Landrum et al. 2014) lead to a significant reduction of Gs signaling under low α-MSH stimulation conditions (significance threshold: false discovery rate (FDR) < 1%; **Fig. 2C**). Variants that are significantly loss-of-function in this condition are rarer in the human population, and more common human variants have no significant effect on MC4R function (significance threshold: FDR < 1%; **Fig.**

2D). Loss-of-function variants by our DMS assay are also typically (e.g., AlphaMissense: 93.4%, 1894/2028) predicted to be deleterious by commonly used variant effect predictors like AlphaMissense (Cheng et al. 2023) and popEVE (Orenbuch et al. 2023) (**Supplementary Fig. 5**).

Because of the sensitivity of our reporter system and the statistical power gained by testing dozens of unique barcodes per variant, we anticipated that these assays would capture subtle quantitative, rather than just qualitative, effects on MC4R function. To assess this, we benchmarked our results against previous quantitative characterizations of MC4R variants from the literature (**Fig. 2E-F**). For example, many MC4R variants that have been observed in the human population have been previously tested for their effects on MC4R function, and a review (Huang, Wang, and Tao 2017) summarizing this work systematically classified >70 variants according to whether they result in “full”, “partial”, or “no/mild” loss of MC4R activity. Our results are consistent with these classifications: variants classified previously as “full” loss-of-function typically have very low MC4R activity in our assay, “partial” variants have intermediate effects, and “no/mild” effect variants have near-normal activity (**Fig. 2E,** median log2[fold change of variant activity over wild-type activity] of -1.4, -0.8, and -0.3, respectively, for each group). Finally, our results show a high degree of correlation (Pearson correlation = 0.84, R^2 = 0.71) with quantitative effect measurements reported for 25 variants individually introduced into the orthosteric site of MC4R (Zhang et al. 2021). Collectively, this demonstrates that our high quality MC4R DMS data accurately and quantitatively assess the effects of variants on MC4R’s function.

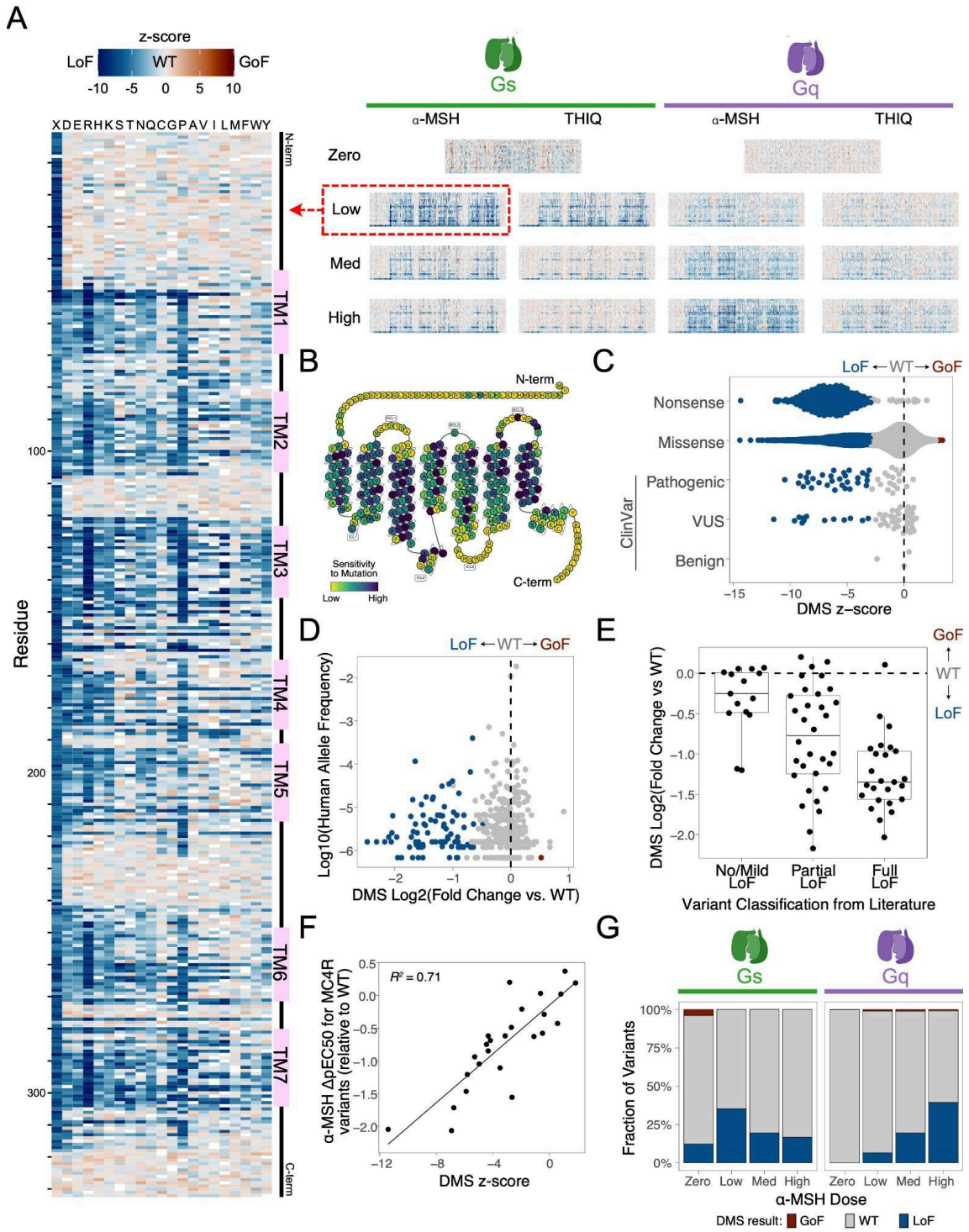


Figure 2. Effect of 6,633 missense and nonsense variants on MC4R signaling functions. A. Heatmaps showing the functional effects (z-scores) for nearly all possible amino acid substitutions (6,633 of 6,640) on MC4R activity for two GPCR signaling functions (Gs and Gq) under a variety of conditions. Heatmaps showing results (both z-score and log2[fold change of variant activity over wild-type]) for all experimental conditions are shown in **Supplementary Figs. 3-4**. The results of the Gs assay with low α-MSH stimulation are highlighted on the left (and in Panels **B-F**). TM: transmembrane domain; GoF: gain-of-function; LoF: loss-of-function; WT: wild-type activity **B.** A modified snake plot (adapted from https://gpcrdb.org/) showing the sensitivity of each MC4R residue to mutation (defined as the mean

log2[fold change variant activity over wild-type] divided by sqrt(sum(standard error^2)) after excluding nonsense variants). **C.** Z-scores for each variant (point), broken out by variant type and clinical (ClinVar) classification. Blue indicates statistically significant LoF, brown is significant GoF (significance threshold: FDR < 1%). VUS: variant of uncertain significance. **D.** Functional effect (Log2[fold change of variant activity over wild-type], x-axis) for all human variants relative to the frequency of the allele in the in the human population (y-axis, gnomAD global population (S. Chen et al. 2024)). **E.** DMS results for human MC4R variants (y-axis) relative to previous functional classifications in the literature (x-axis, from (Huang, Wang, and Tao 2017)) **F.** DMS results (z-scores, x-axis) compared to change in α-MSH potency (relative to WT) of 25 MC4R variants made to the orthosteric binding site for α-MSH (y-axis, cAMP accumulation assay results from (Zhang et al. 2021)). **G.** Fractions of MC4R variants that result in LoF, GoF, or WT activity for eight unique experimental conditions.

Systematic human variant interpretation of MC4R

Looking across multiple data sets gives a comprehensive picture of variant effects, as the various experimental conditions tested have disparate strengths and weaknesses at detecting loss-of-function (LoF) versus gain-of-function (GoF) activities (**Fig. 2G**). For example, unstimulated conditions (e.g., zero α-MSH) uncover variants that lead to constitutive activation of MC4R, but they have less ability to detect loss-of-function variants. In contrast, conditions with agonist stimulation are much better enabled to identify loss of MC4R function. Considering *individual* α-MSH stimulation conditions (zero, low, medium, and high) for both the Gs and Gq assays, each condition identifies 6.6 - 39.3% of variants as loss-of-function and 0.02 - 1.1% as gain-of-function (**Fig. 2G**). Collectively across all α-MSH stimulation conditions, 3,370 variants (50.8%) are loss-of-function in at least one condition, 347 variants (5.2%) are gain-of-function in at least one condition, and 2,996 (45.2%) always show wild-type activity. Interestingly, 80 (1.2%) variants are classified as both loss- and gain-of-function, depending upon the condition.

To aid in clinical variant interpretation, we provide detailed functional effect classifications for 220 human variants reported in ClinVar (Landrum et al. 2014) or from published patient sequencing studies (Farooqi et al. 2003; Yeo et al. 2003; Hinney et al. 2006; Stutzmann et al. 2008; Wade et al. 2021; Brouwers et al. 2021; Huang, Wang, and Tao 2017; Hinney, Volckmar, and Knoll 2013; Rodríguez Rondón et al. 2024) in **Supplementary Table 2**. In total, 130 of these reported human variants (59.1%) are LoF in at least one condition, consistent with being pathogenic for obesity-related phenotypes. This includes 83.9% (26/31) of the variants that are

reported in ClinVar as pathogenic or likely pathogenic, 53.3% (32/60) of those that are unclassified or have conflicting interpretation, and 0% (0/3) of those classified as benign or likely benign. A small number of reported human variants (V103I, H158R, I251L) result in significant increases in Gs and/or Gq signaling, consistent with having a protective effect for obesity, and are generally classified as benign in ClinVar or as having wild-type activity in the literature (Hinney et al. 2006). These results highlight the utility of systematic deep scans across multiple experimental conditions for facilitating human variant interpretation.

Variants that bias MC4R function

MC4R signals through multiple G protein pathways (Sharma et al. 2019; Paisdzior et al. 2020; Ju, Cho, and Sohn 2018; Breit et al. 2011), and evidence from human variant interpretation studies (Lotta et al. 2019; Metzger et al. 2024; Rodríguez Rondón et al. 2024), mouse modeling (Y.-Q. Li et al. 2016), and previous drug programs (Sharma et al. 2019; Clément et al. 2018) suggest that biasing MC4R's activity toward or away from specific pathways could be therapeutically valuable for the treatment of obesity. To gain a better understanding of how MC4R structure relates to its various functions, we systematically searched for variants that differentially impact Gs versus Gq signaling. We applied Principal Component Analysis (PCA) across eight total DMS datasets (four each for Gs and Gq: with zero, low, medium, and high α-MSH stimulation; see **Supplementary Table 1** for details). The first two principal components explained 66% and 12% of the variance, respectively (**Fig. 3A, Supplementary Fig. 6A,B**). Through inspection, we found that Principal Component 1 (PC1) separates variants that impact both signaling functions, with variants that are loss-of-function for both Gs and Gq having higher PC1 values (**Supplementary Fig. 6A,B**). PC2 separates variants that affect Gs and Gq signaling differently (**Fig. 3A-D, Supplementary Fig. 6A,B**). Variants with higher PC2 values typically exhibit Gq bias by having greater than wild-type levels of Gq signaling activity, while retaining wild-type levels of Gs activity (**Fig. 3C,D; Supplementary Fig. 6C,D**). In contrast,

variants with more negative PC2 values are Gs-biased variants that typically have wild-type levels of Gs signaling and reduced Gq signaling upon agonist stimulation (**Fig. 3C**).

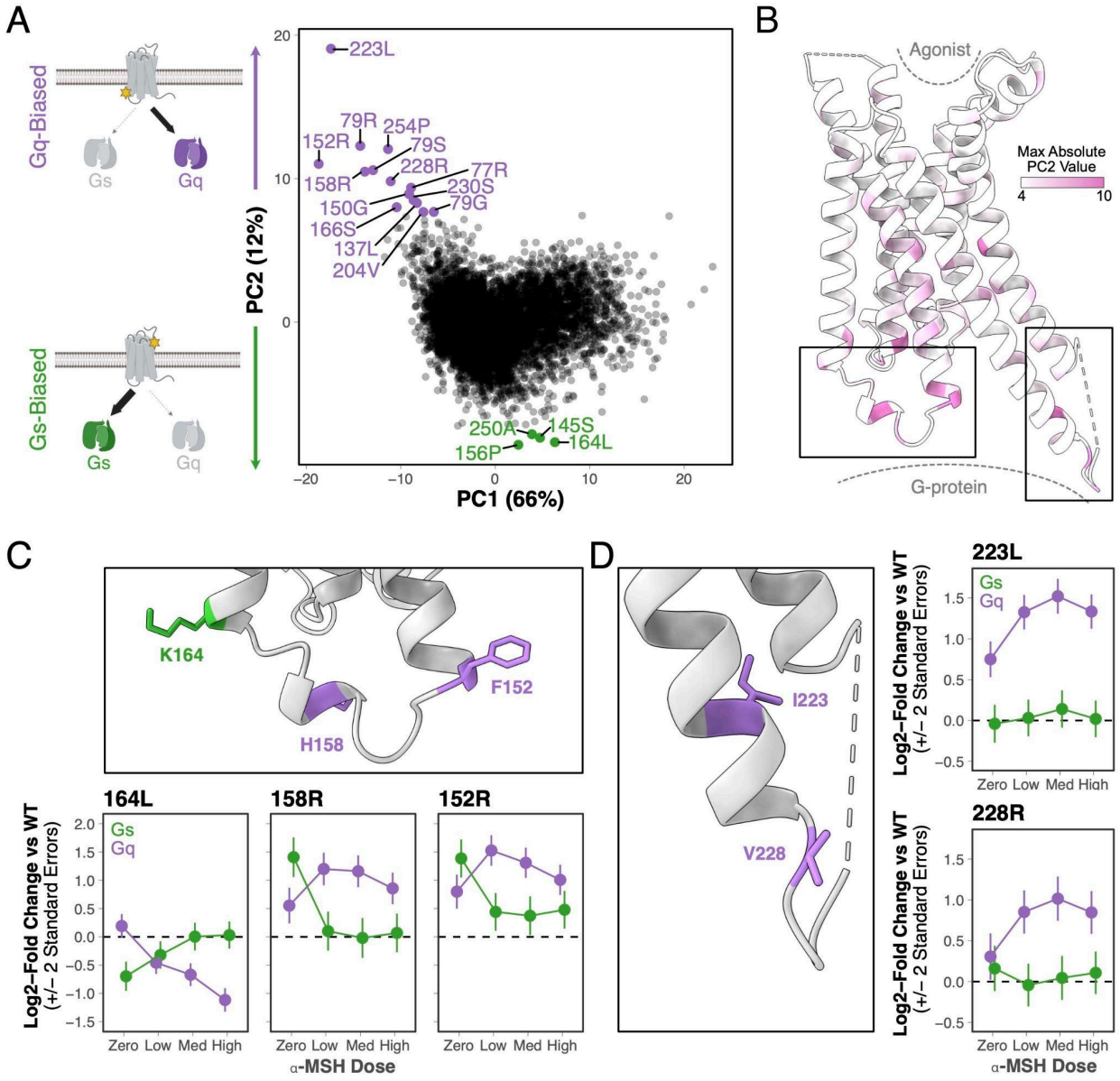


Figure 3. Systematic identification of variants that have biased effects on MC4R signaling. A. Principal component analysis of eight MC4R DMS conditions (Gs and Gq signaling each at four α-MSH stimulation doses: zero, low, medium, high, see **Supplementary Table 1**). Each point indicates a unique variant, with those exhibiting extreme Gq or Gs bias colored in purple and green, respectively. PC, principal component. **B.** Ribbon structure of MC4R showing the maximum absolute PC2 value for each residue position. **C. Top:** Closeup of MC4R structure shown in bottom left box in **B**, highlighting positions where variants result in Gs (green) or Gq (purple) bias. **Bottom:** MC4R signaling activity (log2[fold change of variant activity relative to wild-type]) for three selected variants across α-MSH doses. Error bars are +/- 2 standard errors. Med: medium; WT: wild-type **D. Left:** Closeup of MC4R structure shown in bottom right box in **B**, highlighting positions where variants result in Gs or Gq bias as in **C**. **Right:** MC4R signaling activity for select variants as in **C**.

Overall, more variants substantially increase Gq signaling than Gs signaling (**Fig. 3A**). Gs is the primary G protein coupling for MC4R (Tao 2010; Podyma et al. 2018), and our data suggest that there is little room for further improving MC4R's already robust Gs signaling activity. Interestingly, biased variants are positionally diverse. For example, the 14 variants that display the most extreme Gq bias (**Fig. 3A**) are found at 12 different residue positions, with position 79 unique in having several variants that result in Gq bias (**Supplementary Fig. 6D**). Many of the variants with extreme Gq or Gs bias are located within the regions of MC4R that interact with G proteins, with some scattered throughout the transmembrane domains and far fewer in the vicinity of the protein's orthosteric site (**Fig. 3B-D**).

Structural insights into biased signaling

Comparing these results with existing structural information could provide additional detailed insights into MC4R's signaling functions. It is thought that ligand binding in GPCR orthosteric sites is communicated to the intra-cellular G-protein binding domain through a series of conserved residues or "microswitches" (Hauser et al. 2021; Zhou et al. 2019). Structural studies comparing the inactive and active state structures have confirmed that MC4R shares a similar signaling cascade (Yu et al. 2020; Zhang et al. 2021). Upon ligand binding, W258 (W258^{6x48} in https://gpcrdb.org/ nomenclature, where 6 corresponds to the 6th transmembrane helix and 48 denotes 258 is 2 residues before the most conserved residue in that helix (Isberg et al. 2015)) of the conserved CWxP motif undergoes a conformational rearrangement that is translated to L133^{3x36} and I137^{3x40}, of the conserved PIF motif (MIF in melanocortin receptors). This causes F254^{6x44} in the PIF motif to rearrange, which in turn disrupts the packing of three different interactions: 1) L140^{3x43} and I143^{3x46}, 2) I251^{6x41} and L247^{6x37}, and 3) R147^{3x50} and N240^{6x30}. These, amongst other rearrangements, culminate in the receptor being able to bind a G-protein. This interaction with the G-protein is primarily mediated through R147^{3x50} in the conserved DRY motif, T150^{3x53}, Y157^{34x53}, H158^{34x54} and R305^{7x56} (Zhang et al. 2021).

Strikingly, a number of mutations at residues throughout this signaling cascade had extremely positive PC2 values, implicating them as Gq-biasing mutations. Within the core of the cascade, we identified I137^{3x40}L, F254^{6x44}P, and L140^{3x43}I as Gq-biased (**Fig. 3A; Supplementary Fig. 6C**). We also identified a number of Gq-biasing mutations within the G-protein binding pocket, specifically T150^{3x53}G and H158^{34x54}R (**Fig. 3A,C; Supplementary Fig. 6D**). Interestingly, the H158^{34x54}R variant is found in the human population (Hinney et al. 2006; Wade et al. 2021) and has previously been shown to preferentially signal through the Gq pathway (Paisdzior et al. 2020). H158^{34x54} is also co-located near two other mutations (K164^{3x55}L, F152^{3x55}R), in intracellular loop 2 (ICL2) and near the ends of the third and fourth transmembrane domains (TM3 and TM4), that display bias (**Fig. 3C**). Interestingly, K164^{3x55}L exhibits a Gs bias in that it drives loss of function through Gq.

Our data also point to a number of potentially novel interactions. For example, M79^{2x39} packs against residues H387$^{G.H5.19}$ and Q390$^{G.H5.22}$ of the Gs alpha subunit ($G_s\alpha$) (Flock et al. 2015). This position has multiple different variants that result in Gq bias, including M79^{2x39}R, M79^{2x39}S, and M79^{2x39}G (**Fig. 3A, Supplementary Fig. 6D**). The most extreme bias signal in our PCA analysis came from I223^{5x69}L (**Fig. 3A,D**), which interfaces with the C-terminus of $G_s\alpha$ (Flock et al. 2015), near position L394. Further down towards the intracellular side of TM5, we also identified V228R (**Fig. 3A,D**), which interfaces proximal to E323$^{G.hgh4.13}$ in $G_s\alpha$ (Flock et al. 2015). Collectively, the combination of DMS data and structural information is a fruitful avenue for generating protein structure-function hypotheses. These results highlight the power of DMS to identify regions of MC4R that could be harnessed for designing drugs that precisely modulate specific cellular signaling functions.

Systematic prediction of treatment response

Many variants of MC4R disrupt signaling by causing protein misfolding, which ultimately inhibits proper localization of MC4R to the cell membrane (Huang, Wang, and Tao 2017). Correctors are a class of small molecule drugs that facilitate protein folding. Corrector therapies have been developed for phenotypes such as cystic fibrosis (Boyle and De Boeck 2013) and Fabry disease (Germain et al. 2016), and they have been proposed as a strategy for treating *MC4R*-associated obesity (Huang, Wang, and Tao 2017; Wang et al. 2014; Huang and Tao 2014). One feature of corrector therapies is that they are typically only effective, and therefore FDA-approved, for a subset of patients harboring specific sequence variants (Boyle and De Boeck 2013; Weaver et al. 2022). Identifying variants that respond to corrector therapy is typically done by rigorously testing the effect of a compound on a single variant at a time (Weaver et al. 2022), and DMS offers an attractive avenue to systematically test the treatment response of thousands of patient variants in a single assay.

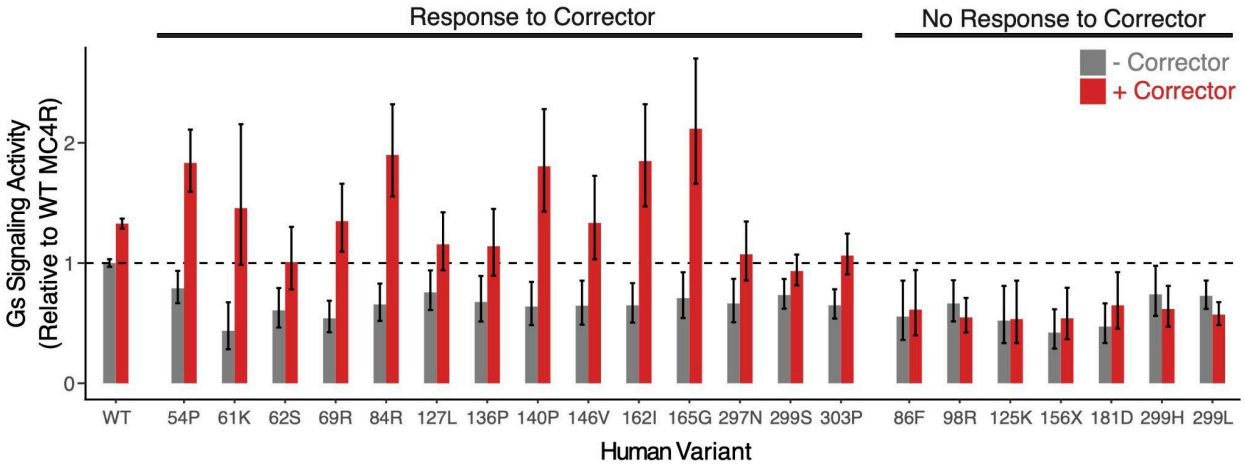


Figure 4. Corrector therapy rescues the activity of a subset of human MC4R variants. Gs signaling activity of 21 selected MC4R variant alleles (of 6,633 tested) with (red) and without (gray) Ipsen 17, a small molecule corrector that has been shown to restore the activity of misfolded MC4R. Bars represent the activity of the variant allele normalized to that of WT MC4R in the no corrector condition, and error bars are +/- two standard errors.

To this end, we tested whether Ipsen 17, a small molecule tool compound that has been shown to correct MC4R misfolding (Wang et al. 2014; Poitout et al. 2007), is able to restore the Gs signaling function of the MC4R variants in our DMS library. Out of all 6,633 tested variants, 290 (4.4%) showed disrupted Gs signaling in the absence of treatment that was partially or fully rescued by the addition of Ipsen 17 (**Supplementary Fig. 7**, see **Methods** for details of statistical analysis). This includes a number of variants that have been classified as pathogenic in ClinVar (Landrum et al. 2014) or otherwise found in patient sequencing studies (Farooqi et al. 2003; Yeo et al. 2003; Hinney et al. 2006; Stutzmann et al. 2008; Wade et al. 2021; Brouwers et al. 2021; Huang, Wang, and Tao 2017; Hinney, Volckmar, and Knoll 2013) (**Fig. 4** shows results for selected variants reported in the human population). Other reported patient variants showed no functional improvement in response to corrector therapy (**Fig. 4**). Collectively, these data support that performing DMS in the presence of a small molecule corrector can be used to systematically predict which patients are likely to benefit from such treatment modalities.

Mapping protein-ligand interactions

DMS experiments can be used to define “drug-resistant” variants within MC4R that disrupt the activity of different types of ligands, providing functional insight into protein-ligand interactions that are key for understanding the mechanisms underlying agonism. Such functional information would be a valuable addition to structural methods and has the potential to streamline the lengthy and iterative cycle of compound optimization in drug discovery. Substantial work has been done to characterize how peptide agonists interact structurally with MC4R, but similar work on small-molecule agonists with comparable activity and selectivity remains relatively limited (Gonçalves, Palmer, and Meldal 2018; Sharma et al. 2019; Yu et al. 2020; Heyder et al. 2021; Zhang et al. 2021). To characterize the functional interaction landscapes of different ligand types and to better understand what distinguishes peptide and small-molecule MC4R pharmacophores, we performed DMS assays of MC4R using both native peptide agonist

stimulation (α-MSH) and small molecule agonist stimulation (THIQ). Bayesian meta-regression analysis of the lowest dose concentrations of α-MSH and THIQ for the Gs signaling assay revealed a set of variants that uniquely disrupt activation by one ligand but not the other (at FDR < 5%, **Fig. 5A**). These variants cluster exclusively within the orthosteric binding pocket (**Fig. 5B-E**) and at positions of known binding interactions of each molecule (Zhang et al. 2021). Notably, there are many more variants that uniquely disrupt activation by α-MSH (**Fig. 5A-D**). For example, a majority of amino acid substitutions at position I104^{2x6} disrupt activation of MC4R by α-MSH, but none lead to significant reduction in MC4R activity upon stimulation by THIQ (**Fig. 5C**). Multiple positions within the orthosteric binding pocket displayed this pattern (**Fig. 5B-E**), which is consistent with how the peptide agonist utilizes a larger network of interactions to increase binding affinity.

Interestingly, this comparison also highlights how the same residue of MC4R can be critical for interfacing with multiple ligands but points to substantive differences in the precise physical interaction between each ligand and that position of the target. For example, positions P48^{1x36} and I129^{3x32} harbor variants that can have differentially deleterious effects under the two ligand conditions (**Fig. 5C-E**). Many amino acid substitutions at P48^{1x36} (including to V, I, F, Y, Q, and R) disrupt stimulation of MC4R by α-MSH but not THIQ. However, changing this position to the negatively charged P48^{1x36}D variant has the opposite effect, uniquely ablating activation by THIQ. The HFRW motif (His6-Phe7-Arg8-Trp9) of α-MSH represents a conserved pharmacophore critical for activation of MC4R by peptides, and the tri-branched THIQ molecule (R1-R2-R3) mimics the HFRW conformational architecture (**Fig. 5E**) (Hruby et al. 1987; Gonçalves, Palmer, and Meldal 2018; Zhang et al. 2021). Our observation that α-MSH stimulation is more sensitive to variants at position P48^{1x36} is consistent with how the His6 of α-MSH forms more interactions with this hydrophobic pocket of MC4R, while the analogous R3 group of THIQ is more flexible and forms non-specific interactions in this region (**Fig. 5C–E**)

(Hruby et al. 1987; Gonçalves, Palmer, and Meldal 2018; Zhang et al. 2021). At position I129^{3x32}, mutation to any polar uncharged variant (S, T, N, or Q) alters THIQ activation, while I129^{3x23}V uniquely inhibits α-MSH activation (**Fig. 5C–E**). The Phe7 of α-MSH and the analogous R2 group of THIQ form key interactions with Ca^{2+} and the core hydrophobic pocket formed by I129^{3x32} of MC4R (**Fig. 5E**) (Zhang et al. 2021). The enrichment of variants at I129^{3x32} that uniquely disrupt MC4R activation by THIQ points to a stronger dependency of the small-molecule on interactions with this residue. In summary, by assessing the functional consequences of all mutations within the MC4R orthosteric site, we not only confirm known binding interactions but also reveal interactions that distinguish peptide and small-molecule activation. These relationships provide additional functional insight into the structural mechanism of MC4R ligand binding that could be harnessed for drug design.

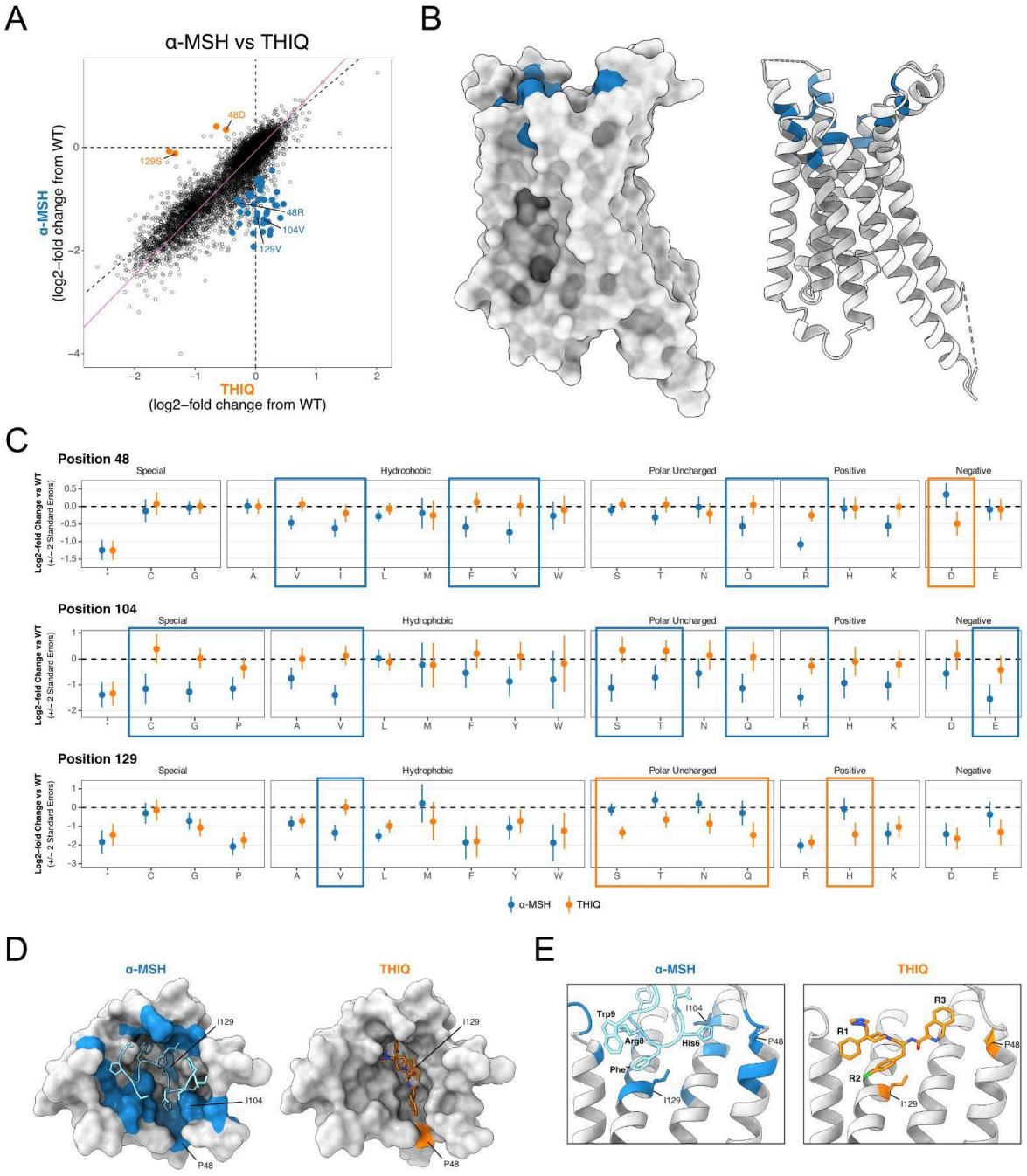


Figure 5. Systematic identification of functional protein-ligand interactions. A. Bayesian meta-regression of the α-MSH and THIQ datasets at low agonist concentration for Gs reveals variants that differentially affect MC4R activation by each agonist. Statistically significant effects are colored by which agonist condition they impaired (α-MSH, blue; THIQ, orange, FDR < 0.05). Variants at positions 48, 104, and 129 are labeled. **B.** Side-views of MC4R structure in surface (left) and ribbon (right) view show clustering of variants from **A** that specifically abrogate α-MSH but not THIQ activation (blue) in the extracellular orthosteric binding site. **C.** Effect of all possible variants at residues 48, 104, and 129 at low concentration of α-MSH (blue) and THIQ (orange). Variants that disproportionately affect activation by only one agonist are boxed by respective color. **D.** Top-down surface views of MC4R (PDB: 7f58) with bound α-MSH (left, blue; PDB: 7f53) or THIQ (right, orange; PDB: 7f58). Positions are colored by whether variants at that position uniquely perturb activation by α-MSH (blue) or THIQ (orange). **E.** Zoomed view of the MC4R binding pocket with α-MSH bound (left) or THIQ bound (right). Structures are colored as in (D). Residues that form the HFRW motif of α-MSH and functional groups R1, R2, and R3 of THIQ are labeled in bold. Residues 48, 104, and 129 are shown in stick form.

Discussion

Deep Mutational Scanning holds substantial potential for improving many phases of drug discovery and development, but realizing this potential requires the ability to draw highly quantitative and disease-relevant conclusions from DMS data, an outstanding challenge for the field. Here, we demonstrated the value of improved DMS assays and analysis methods in addressing these challenges by applying our approaches to a medically relevant GPCR, MC4R. First, we were able to quantify differences in Gs-mediated cAMP and Gq-mediated calcium signaling for MC4R in response to its native ligand, α-MSH. Understanding these subtle differences in molecular signaling phenotypes is crucial for driving precision treatments. For example, Gq signaling bias is one of the purported mechanisms for the better side-effect profile of the FDA approved MC4R agonist, Setmelanotide (Sharma et al. 2019). Additionally, we systematically assessed the responses of MC4R variants implicated in human obesity to a small-molecule corrector, Ipsen 17. Whereas traditional approaches would have required the separate experimental assessment of every variant, DMS tests every conceivable single amino acid variant in one experiment. This enables researchers to understand how a potential treatment will work for broader swaths of the population before drugs reach the clinic, which could ultimately lead to more effective clinical trial designs.

Our data also contribute to the growing understanding of the complicated mechanism by which MC4R (and other GPCRs) translate ligand binding on the extracellular side of the membrane, through an allosteric network of residues, to G-protein activation on the intracellular side (Howard et al. 2024). We demonstrate the utility of DMS for understanding how proteins bind other molecules. In particular, we identified mutations that uniquely disrupt the binding of two agonists, one a peptide (α-MSH) and one a small molecule (THIQ). This is crucial information for understanding the mechanism by which these molecules bind MC4R, which could lead to hypotheses for further optimization of more potent compounds. DMS data contain a substantial

amount of latent information about protein structure, and can help identify novel pockets to target with new compounds (Weng et al. 2024). As suggested above, we also envision using DMS data to help optimize molecules by identifying mutations that uniquely disrupt or potentiate sets of compounds in a chemical series.

Importantly, the experimental and analysis frameworks we describe are widely applicable to GPCRs and more broadly to other target classes. GPCRs are known to signal through a variety of pathways, and there are existing transcriptional reporters for most G-protein mediated signaling pathways that could be harnessed to gain a holistic understanding of the downstream consequences of GPCR activity (Hauser et al. 2022). Furthermore, there are a substantial number of existing transcriptional reporters for a wide variety of cellular signaling processes mediated by transcription factors, kinases, and other major targets of drug discovery efforts, which could extend these methods to other target classes.

Finally, there has been incredible progress over the last few years in the development of machine learning-based models to predict the structural and functional consequences of mutations on proteins. That said, these models have struggled to address many real-world applications in the treatment of human disease (McDonald et al. 2024). One potential explanation to this gap in performance could be due to the fact that many DMS datasets (our previous work included) used to train these models typically assay a limited set of experimental conditions, often assay cellular effects that are only indirectly related to disease etiology, suffer from low signal to noise ratios, have no quantification of uncertainty in their measurements, and are subject to a heterogeneous set of subtle experimental caveats. Detailed experimental characterization of protein function (such as the >21,500,000 different measurements in this work alone) and efforts such as ProteinGym (Notin et al. 2023) to benchmark large-scale functional data will continue to be critical for developing and evaluating large scale models that

can be applied to a wider variety of drug discovery and development applications. As both the experimental and computational methods continue to improve, we foresee DMS having a profound impact on the drug discovery process.

Materials and Methods

Reporter assays and cell line development

The Gs (cAMP-luciferase) reporter assay, diagramed in **Supplementary Fig. 1A**, was adapted from an assay previously used to assess the function of other GPCRs (Jones et al. 2020, 2019).

The Gq assay relay reporter system (diagrammed in **Supplementary Fig. 1B**) is described in detail in a patent stemming from this work (Chan, Cooper, and Gasperini 2024). Briefly, it was constructed as follows: A piggybac transposon plasmid was constructed using Gibson assembly harboring a genetic cassette that expresses a synthetic transcription factor, Gal4_DBD-VPR under the control of an NFAT response element. Human embryonic kidney cells (HEK293T) were then cotransfected with this plasmid along with a piggybac transposase expression vector, and cells were selected for puromycin resistance. These cells were then isocloned by dilution plating, and colonies were selected based on single cell clones. Five clones were then carried forward based on cell morphology and growth rates. A second series of Bxb1-based landing pad vectors were constructed containing a library of 20 Gq-coupled GPCRs under the control of the dox-inducible promoter, along with a second genetic cassette containing a DNA-barcoded luciferase gene under the control of the Gal4_UAS enhancer. This plasmid library was integrated into each of the five isoclonal Gq-relay cell lines. The best performing Gq-relay cell line was selected on signal-to-noise criteria across a panel of agonists for the 20 Gq-coupled GPCRs. Use of this relay system in combination specifically with an *MC4R* transgene resulted in α-MSH-dose dependent expression of the reporter gene (**Supplementary Fig. 1C**).

Building the DMS libraries

Generation and genomic integration of the plasmid libraries containing all possible single amino acid substitutions of MC4R used a further optimized version of an earlier method (Jones et al. 2020). In brief, variant segments of *MC4R* cDNA were amplified from DNA microarrays and

cloned into base vectors through a multi-step process to yield pooled libraries of *MC4R* variants with fully intact and barcoded reporter gene cassettes. Fully assembled plasmid libraries were then co-transfected with a plasmid encoding Bxb1 recombinase into HEK293T cells containing a landing pad at the H11 safe harbor locus to achieve single copy integration per cell. In a deviation from the previously published method, two independent replicates of each sub-library were cloned and pooled together post-cellular integration in order to maximize library coverage and the number of barcodes per variant.

Variant-barcode mapping

As described above, barcodes were randomly appended to variants during amplification of MC4R segments from variant oligo pools and then ligated into library base vectors. After this first step of plasmid library cloning, variant segments were amplified along with the neighboring barcode and sequenced with 2x150 paired-end reads (see **Supplementary Table 3** for amplification and sequencing primers) on an Illumina NextSeq 550 instrument using a 300-cycle Mid Output kit. Illumina 2x150 BCL files were demultiplexed with *bcl2fastq2* into R1 and R2 FASTQ files, which were merged into single fragments using *Flash2* requiring a minimum 5 bp of overlap (Quan et al. 2019). The first 21 bp of each fragment corresponding to the barcode sequence were extracted into the read name, and the remaining fragment was adapter-trimmed using *umi_tools* and *cutadapt*, respectively (Smith, Heger, and Sudbery 2017; Martin 2011). The remaining fragments were mapped against a custom reference composed of the designed oligonucleotide library using *STAR* with default parameters except requiring that alignments be strictly unique to be reported (Dobin et al. 2013). Taking each alignment as an oligo-barcode pair, the read counts per unique oligo-barcode pair were computed for each replicate and joined by barcode. Finally, the resulting maps were filtered to require each oligo-barcode pair to pass three requirements in both replicates: correct barcode length, total read depth > 10, and purity > 0.75. The purity of an oligo-barcode pair was defined as the read count of that pair divided by

the total number of reads containing that barcode. Post-processing after *STAR* was performed using *samtools* for BAM manipulation and custom R code (H. Li et al. 2009).

Running DMS assays

The DMS assay protocol was optimized from a previous method (Jones et al. 2020). HEK293T single-copy variant cell libraries were seeded at a density of ~$17x10^6$ cells per 150 mm tissue-culture treated dish in DMEM + 10% fetal bovine serum (FBS). Four dishes were seeded for each experimental condition, with each dish being treated as an independent biological replicate (4 replicates per condition). Twenty-four hours after seeding, media was exchanged with DMEM + 0.5% FBS +/- 10 ng/mL Doxycycline. For chaperone experiments, all conditions were additionally replicated +/- 1 μM Ipsen 17. 24 hrs after Doxycycline induction, media was exchanged with Opti-MEM + DMSO, Forskolin, or MC4R agonist (α-MSH or THIQ). Forskolin bypasses MC4R to constitutively activate cAMP signaling, so this condition was used as a variant-independent measurement of library composition. For chaperone experiments, cells were washed 3x with 10 mL DMEM to remove Ipsen 17 prior to agonist stimulation as it has been shown to be an antagonist of α-MSH activity and is thought to bind directly to the same site on MC4R (Wang et al. 2014). Six hours after agonist stimulation, cells were harvested by scraping in 4 mL lysis buffer (RLT buffer (Qiagen) + 143 mM β-ME). Given the seeding density (~$17x10^6$ cells per 150 mm replicate dish), time from seeding to collection, and doubling time of HEK293T cells, approximately $25.5x10^6$ cells were collected per replicate. This translates to approximately 30-60x cellular coverage per amino acid variant in each replicate. Lysis was performed by passing the cell slurry 6x through a sterile 18G needle and then spinning through QIAshredder (Qiagen) columns. RNA was extracted from 1 ml of the homogenized lysate with the RNeasy Plus Mini kit (Qiagen), including optional on-column DNAse digestion, and eluted into 100 μL water. Eight reverse transcriptase reactions per sample were performed with the SuperScript IV kit (Thermo Fisher), as described previously (Jones et al. 2020) (primers listed in

Supplementary Table 3). cDNA from each sample was treated with 1 µL RNase A (100 µg/ml, Thermo Fisher) and 3.2 µL RNase H (5,000 U/mL, NEB) at 37°C for 30 min. RNase-treated samples were concentrated to ~55 µL by spinning through Amicon Ultra 10 kDa concentrators (EMD Millipore) for ~8 min. To determine the necessary cycle numbers for equivalent amplification of each sample library, qPCR reactions were performed on 1 µL cDNA (diluted 1:8 in water) with Q5 polymerase (NEB), SYBR Green (Thermo Fisher), and library amplification primers (**Supplementary Table 3**). Final amplification cycles for each sample were chosen by adding three cycles to the Cq values generated from each respective qPCR reaction. Illumina sequencing libraries were prepared by amplifying 50 µL of each cDNA sample with sequencing adapters (500 nM each library amplification primer, **Supplementary Table 3**) using the NEBNext Q5 High Fidelity 2x PCR Kit (NEB) under the following cycling conditions: 98°C for 30 s, X cycles of 98°C for 8 s, 65°C for 20 s, and 72°C for 10 s, followed by an extension of 72°C for 2 min. 3 µL of each DNA library sample was run on a 4% E-Gel (Thermo Fisher) and densitometry was performed with Fiji to account for differences in library yields. Samples were mixed at equal amounts into a single pool and then purified into 200 µL IDTE (Qiagen) with AxyPrep Magnetic beads (Fisher Scientific). The purified library was quantified with the DeNovix High-Sensitivity Fluorescence kit and prepared for sequencing with a 10% PhiX spike-in. Final library mixture was sequenced using custom read and index primers (**Supplementary Table 3**) on an Illumina NextSeq 550 with the High Output 75 cycle kit.

Sequence processing for barcode expression

Illumina 1x26 BCL files were demultiplexed with *bcl2fastq2* and processed to remove the last 5 bp using basic bash commands. The resulting sequences were counted for each sample, and the resulting barcodes were joined with the appropriate oligo-barcode map. The resulting barcodes were joined with sample and MC4R variant metadata and returned for regression

analysis. All processing after demultiplexing was performed with custom R code. Total mapped reads per replicate at the RNA-seq stage were as follows:

- Gs/CRE: 9.1-18.2 million mapped reads, median=12.3
- Gq/UAS: 8.6-24.1 million mapped reads, median=14.5
- Gs/CRE+Chaperone: 6.4-9.5 million mapped reads, median=7.5

The median read counts per sample per barcode were 8, 10, and 6 reads for Gs/CRE, Gq/UAS, and Gs/CRE+Chaperone assays, respectively. The median number of barcodes per variant across all samples (the “median of medians”) were 56 for Gs/CRE, 28 for Gq/UAS, and 44 for Gs/CRE+Chaperone. The correlation (r) of barcode readcounts between replicates was ~0.5 and ~0.4 for the Gs and Gq assays, respectively (**Supplementary Fig. 1E**).

Negative binomial regression analysis pipeline

In many DMS methods, barcodes are summed within protein-coding variants to generate a single variant-level count per sample. This unnecessarily sacrifices available power obtained by repeatedly measuring the same molecular process in distinct cells. An alternative is to model barcode counts directly. However, to apply standard models, missing data must either be removed, which can remove a majority of detected barcodes, or somehow imputed. Additionally, many existing methods either use a log transformation of read counts to obtain approximate normality, or apply Poisson regression and related assumptions for inference (Rubin et al. 2017; Faure et al. 2020). However, there is strong prior reason to believe that DMS counts are overdispersed in our data, since our synthetic reporter system is read out via RNA-seq (Robinson and Smyth 2007).

Instead, we developed and applied a mixed effects negative binomial general linear model (GLM) to resolve these challenges. These models have been widely deployed to model count data, and in particular identify differential expression, in bulk and single-cell RNA-seq analysis.

We implement maximum likelihood estimation for this model using *glmmTMB*, which can accommodate the potentially large scale of multiplex count data (Brooks et al. 2017). For each position, we consider all variants located at that position along with all wild-type variants in the same protein subregion and apply the following model:

$$count_{ijkm} \sim NB(\mu_{ij}, \theta)$$
$$\alpha_k \sim N(0, \sigma^2)$$
$$log(\mu_{ij}) = \beta_i^{condition} + \beta_{ij}^{variant} + \alpha_k + offset(\gamma_m)$$

For the *i*th condition, the *j*th variant, the *k*th barcode, and the *m*th sample. Consequently, the first two terms in the last equation above correspond to a global mean term for each condition and a term for the variant-specific deviation from wild-type in each condition. The last two terms are the random effect for barcode *k*, and the sample-specific technical offset for sample *m*. The definition of the offset is often context specific, and here we use the log of the sum of barcode counts derived from stops, reasoning they should be constant across conditions and replicates.

We fit this model for each MC4R position independently and extract coefficients for the additive shift in the mean of each variant relative to wild-type. Using the per-condition summary statistics, we obtained Wald test statistics by dividing the effect size by the standard error and computed p-values against the normal distribution. P-values were adjusted for multiple testing using the Benjamini-Hochberg method and thresholded to 1% or 5% FDR where indicated.

To define more complex null hypotheses like chaperone rescue, we extracted marginal means for each variant under each treatment using the *emmeans* package (Lenth 2024). We define the chaperone rescue contrast as the additive shift of each variant in each treatment condition to the wild-type mean specifically in the untreated condition. Since this quantity is a linear contrast across marginal means, we computed the associated contrast estimates and standard errors,

and tested them for significance using the same approach as the per-condition summary statistics.

Comparison to human genetics data and variant effect predictors

Pathogenicity classifications of *MC4R* missense and nonsense variants were obtained from ClinVar (Landrum et al. 2014) on January 5, 2024, and all available annotations were included in the analysis regardless of ClinVar review status metric. Human population frequency data for *MC4R* missense variants were obtained from gnomAD (S. Chen et al. 2024) on January 8, 2024. A comprehensive list of 220 *MC4R* variants that are of potential clinical relevance (**Supplementary Table 2**) was mined from ClinVar (Landrum et al. 2014), along with papers or review articles describing variants identified in human sequencing studies (Farooqi et al. 2003; Yeo et al. 2003; Hinney et al. 2006; Stutzmann et al. 2008; Wade et al. 2021; Brouwers et al. 2021; Huang, Wang, and Tao 2017; Hinney, Volckmar, and Knoll 2013; Rodríguez Rondón et al. 2024). Effect predictions for *MC4R* missense variants were obtained from the public releases of AlphaMissense (Cheng et al. 2023) and popEVE (Orenbuch et al. 2023).

Identification of functionally biased variants

We used Principal Component Analysis to identify biased variants. Specifically, we used the test statistic (log2 fold change divided by the standard error) for the DMSO and all α-MSH conditions in both the Gq and Gs pathways to create a matrix with conditions (defined as Drug_Concentration) as columns, and each individual mutant (defined as Position_Substitution) as rows. We then passed this matrix to R's *prcomp* function with default parameters, and used the first two principal components to visualize the results. Visual inspection of the loadings via a biplot, along with plotting of various variant types (**Supplementary Fig. 6A,B**) revealed that PC1 separates variants based on their overall effect on MC4R function, and PC2 separates variants

based on differential response through the Gq and Gs pathways. We set a simple cutoff of +-7.5 on PC2 to highlight particularly interesting variants.

Identifying variants that respond to corrector treatment

We identified variants whose activity increased upon treatment with Ipsen 17 using a slight modification of the summary statistics from the general-purpose model described above. For each variant and condition, we compute the log2-scaled marginal mean (averaging across barcodes and replicates) and its associated standard error. Then, for each variant we compute the following two summary statistics, and test whether their estimates are significantly different from zero. First, we define a variants' *defect* as the marginal mean of that variant under DMSO treatment minus the marginal mean of WT under DMSO treatment. This quantifies the existence and severity of the variant's intrinsic effect on MC4R activity. Second, we define that variant's *rescue* as the marginal mean of that variant under Ipsen 17 treatment minus the marginal mean of that variant under DMSO activity. This quantifies the magnitude of the (typically) increase in MC4R activity upon Ipsen 17 treatment for each variant, relative to its DMSO-only baseline. After computing the indicated estimate and errors, we perform significance testing as in the general model, where we define Wald statistics as the estimate divided by the propagated error, compute p-values from the normal distribution, and adjust for multiple testing using the Benjamini-Hochberg procedure.

Identifying critical variants for protein-ligand interactions

We identified sets of mutations that specifically inhibited or potentiated MC4R activation in the presence of ligands. We applied the general DMS model (see *Negative binomial regression analysis pipeline*) and extracted log2-scaled fold changes and standard errors for each variant relative to WT, within each ligand-treated condition. To account for systematic differences between ligands across all variants, we applied Bayesian meta-regression via the *brms* R

package (Bürkner 2017) and regressed the α-MSH summary statistics against THIQ, while including the errors in both quantities via the se() and me() brms functions. Finally, to infer significant α-MSH- or THIQ-specific effects, we extracted the residual of each variant relative to the meta-regression best-fit line and tested whether this residual was significantly non-zero based on the posterior sampling performed with *brms*.

Structural modeling

Molecular visualization of variant effects on MC4R was performed with UCSF ChimeraX (Pettersen et al. 2021). For visualization of functionally biased extremes in MC4R (**Fig. 3B**), the maximum absolute PC2 value for each position was calculated and the define attribute function of ChimeraX was used to color the structure of α-MSH-bound MC4R (PDB: 7F53) by these relative values, ranging from white (no bias) to pink (extreme bias). For all other structural panels related to variant bias, positions of interest are colored binarily by green or purple to indicate Gs-bias or Gq-bias, respectively. For visualizing variant effects on protein-ligand interactions (**Fig. 5**), positions with significant variants identified by meta-regression were colored by whether variants at a given position perturb activation uniquely by α-MSH (blue) or THIQ (orange). Where α-MSH and THIQ structures are shown, the respective α-MSH-bound (PDB: 7F53) and THIQ-bound (PDB: 7F58) cryo-EM models were used for visualization. For the depictions in **Fig. 5B**, α-MSH-bound MC4R was used.

Figure generation

Cartoon diagrams of reporter assay designs in **Fig. 1** and **Supplementary Fig. 1A,B** were created with BioRender.com.

Sequence data and software availability

Raw sequencing data are available from SRA under project accession number PRJNA1161152. All code used for analysis and figure generation are available at https://github.com/octantbio/mc4r-dms.

Competing Interests

CJH, NSA, BAO, EMJ, LYC, HC, JBA, JSB, ARC, SK, DED, and NBL are current employees of Octant, Inc. and/or hold shares or options in the company. EMJ, HC, AC, and SK are listed inventors on patents related to this work. All other authors declare no competing financial interests.

Acknowledgements

We would like to thank Jakob Sture Madsen for his intellectual contributions to the project and his helpful comments on the paper.

References

Ahlmann-Eltze, Constantin, and Wolfgang Huber. 2021. “glmGamPoi: Fitting Gamma-Poisson Generalized Linear Models on Single Cell Count Data.” *Bioinformatics* 36 (24): 5701–2.

Araya, Carlos L., and Douglas M. Fowler. 2011. “Deep Mutational Scanning: Assessing Protein Function on a Massive Scale.” *Trends in Biotechnology* 29 (9): 435–42.

Boss, V., D. J. Talpade, and T. J. Murphy. 1996. “Induction of NFAT-Mediated Transcription by Gq-Coupled Receptors in Lymphoid and Non-Lymphoid Cells.” *The Journal of Biological Chemistry* 271 (18): 10429–32.

Boyle, Michael P., and Kris De Boeck. 2013. “A New Era in the Treatment of Cystic Fibrosis: Correction of the Underlying CFTR Defect.” *The Lancet. Respiratory Medicine* 1 (2): 158–63.

Braberg, Hannes, Ignacia Echeverria, Robyn M. Kaake, Andrej Sali, and Nevan J. Krogan. 2022. “From Systems to Structure - Using Genetic Data to Model Protein Structures.” *Nature Reviews. Genetics* 23 (6): 342–54.

Brandes, Nadav, Grant Goldman, Charlotte H. Wang, Chun Jimmie Ye, and Vasilis Ntranos. 2023. “Genome-Wide Prediction of Disease Variant Effects with a Deep Protein Language Model.” *Nature Genetics* 55 (9): 1512–22.

Breit, Andreas, Thomas R. H. Büch, Ingrid Boekhoff, Hans Jürgen Solinski, Ellen Damm, and Thomas Gudermann. 2011. “Alternative G Protein Coupling and Biased Agonism: New Insights into Melanocortin-4 Receptor Signalling.” *Molecular and Cellular Endocrinology* 331 (2): 232–40.

Brooks, Mollie E., Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Mächler, and Benjamin M. Bolker. 2017. “glmmTMB Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling.” *The R Journal* 9 (2): 378–400.

Brouwers, Bas, Edson Mendes de Oliveira, Maria Marti-Solano, Fabiola B. F. Monteiro, Suli-Anne Laurin, Julia M. Keogh, Elana Henning, et al. 2021. “Human MC4R Variants Affect Endocytosis, Trafficking and Dimerization Revealing Multiple Cellular Mechanisms Involved in Weight Regulation.” *Cell Reports* 34 (12): 108862.

Bürkner, Paul-Christian. 2017. “Brms: An R Package for Bayesian Multilevel Models Using Stan.” *Journal of Statistical Software* 80 (August):1–28.

Chan, Henry, Aaron Cooper, and Molly Jeanette Gasperini. 2024. Systems and Methods for Measuring Cell Signaling Protein Activity. USPTO 20220177897A1. *US Patent*, filed February 1, 2022, and issued April 25, 2024.

Chavez, Alejandro, Jonathan Scheiman, Suhani Vora, Benjamin W. Pruitt, Marcelle Tuttle, Eswar P R Iyer, Shuailiang Lin, et al. 2015. “Highly Efficient Cas9-Mediated Transcriptional Programming.” *Nature Methods* 12 (4): 326–28.

Cheng, Jun, Guido Novati, Joshua Pan, Clare Bycroft, Akvilė Žemgulytė, Taylor Applebaum, Alexander Pritzel, et al. 2023. “Accurate Proteome-Wide Missense Variant Effect Prediction with AlphaMissense.” *Science* 381 (6664): eadg7492.

Chen, Kong Y., Ranganath Muniyappa, Brent S. Abel, Katherine P. Mullins, Pamela Staker, Robert J. Brychta, Xiongce Zhao, et al. 2015. “RM-493, a Melanocortin-4 Receptor (MC4R) Agonist, Increases Resting Energy Expenditure in Obese Individuals.” *The Journal of Clinical Endocrinology and Metabolism* 100 (4): 1639–45.

Chen, Siwei, Laurent C. Francioli, Julia K. Goodrich, Ryan L. Collins, Masahiro Kanai, Qingbo Wang, Jessica Alföldi, et al. 2024. “A Genomic Mutational Constraint Map Using Variation in 76,156 Human Genomes.” *Nature* 625 (7993): 92–100.

Clément, Karine, Erica van den Akker, Jesús Argente, Allison Bahm, Wendy K. Chung, Hillori Connors, Kathleen De Waele, et al. 2020. “Efficacy and Safety of Setmelanotide, an MC4R

Agonist, in Individuals with Severe Obesity due to LEPR or POMC Deficiency: Single-Arm, Open-Label, Multicentre, Phase 3 Trials." *The Lancet. Diabetes & Endocrinology* 8 (12): 960–70.

Clément, Karine, Heike Biebermann, I. Sadaf Farooqi, Lex Van der Ploeg, Barbara Wolters, Christine Poitou, Lia Puder, et al. 2018. "MC4R Agonism Promotes Durable Weight Loss in Patients with Leptin Receptor Deficiency." *Nature Medicine* 24 (5): 551–55.

Collet, Tinh-Hai, Béatrice Dubern, Jacek Mokrosinski, Hillori Connors, Julia M. Keogh, Edson Mendes de Oliveira, Elana Henning, et al. 2017. "Evaluation of a Melanocortin-4 Receptor (MC4R) Agonist (Setmelanotide) in MC4R Deficiency." *Molecular Metabolism* 6 (10): 1321–29.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21.

Farooqi, I. Sadaf, Julia M. Keogh, Giles S. H. Yeo, Emma J. Lank, Tim Cheetham, and Stephen O'Rahilly. 2003. "Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene." *The New England Journal of Medicine* 348 (12): 1085–95.

Fatima, Munazza Tamkeen, Ikhlak Ahmed, Khalid Adnan Fakhro, and Ammira Sarah Al-Shabeeb Akil. 2022. "Melanocortin-4 Receptor Complexity in Energy Homeostasis,obesity and Drug Development Strategies." *Diabetes, Obesity & Metabolism* 24 (4): 583–98.

Faure, Andre J., Júlia Domingo, Jörn M. Schmiedel, Cristina Hidalgo-Carcedo, Guillaume Diss, and Ben Lehner. 2022. "Mapping the Energetic and Allosteric Landscapes of Protein Binding Domains." *Nature* 604 (7904): 175–83.

Faure, Andre J., Jörn M. Schmiedel, Pablo Baeza-Centurion, and Ben Lehner. 2020. "DiMSum: An Error Model and Pipeline for Analyzing Deep Mutational Scanning Data and Diagnosing Common Experimental Pathologies." *Genome Biology* 21 (1): 207.

Findlay, Gregory M., Evan A. Boyle, Ronald J. Hause, Jason C. Klein, and Jay Shendure. 2014. "Saturation Editing of Genomic Regions by Multiplex Homology-Directed Repair." *Nature* 513 (7516): 120–23.

Flock, Tilman, Charles N. J. Ravarani, Dawei Sun, A. J. Venkatakrishnan, Melis Kayikci, Christopher G. Tate, Dmitry B. Veprintsev, and M. Madan Babu. 2015. "Universal Allosteric Mechanism for Gα Activation by GPCRs." *Nature* 524 (7564): 173–79.

Fowler, Douglas M., and Stanley Fields. 2014. "Deep Mutational Scanning: A New Style of Protein Science." *Nature Methods* 11 (8): 801–7.

Germain, Dominique P., Derralynn A. Hughes, Kathleen Nicholls, Daniel G. Bichet, Roberto Giugliani, William R. Wilcox, Claudio Feliciani, et al. 2016. "Treatment of Fabry's Disease with the Pharmacologic Chaperone Migalastat." *The New England Journal of Medicine* 375 (6): 545–55.

Gonçalves, Juliana Pereira Lopes, Daniel Palmer, and Morten Meldal. 2018. "MC4R Agonists: Structural Overview on Antiobesity Therapeutics." *Trends in Pharmacological Sciences* 39 (4): 402–23.

Greenfield, Jerry R., Jeffrey W. Miller, Julia M. Keogh, Elana Henning, Julie H. Satterwhite, Gregory S. Cameron, Beatrice Astruc, et al. 2009. "Modulation of Blood Pressure by Central Melanocortinergic Pathways." *The New England Journal of Medicine* 360 (1): 44–52.

Hauser, Alexander S., Charlotte Avet, Claire Normand, Arturo Mancini, Asuka Inoue, Michel Bouvier, and David E. Gloriam. 2022. "Common Coupling Map Advances GPCR-G Protein Selectivity." *eLife* 11 (March). https://doi.org/10.7554/eLife.74107.

Hauser, Alexander S., Albert J. Kooistra, Christian Munk, Franziska M. Heydenreich, Dmitry B. Veprintsev, Michel Bouvier, M. Madan Babu, and David E. Gloriam. 2021. "GPCR Activation Mechanisms across Classes and Macro/microscales." *Nature Structural & Molecular*

Biology 28 (11): 879–88.
Heyder, Nicolas A., Gunnar Kleinau, David Speck, Andrea Schmidt, Sarah Paisdzior, Michal Szczepek, Brian Bauer, et al. 2021. “Structures of Active Melanocortin-4 Receptor-Gs-Protein Complexes with NDP-α-MSH and Setmelanotide.” *Cell Research* 31 (11): 1176–89.
Hinney, Anke, Thomas Bettecken, Patrick Tarnow, Harald Brumm, Kathrin Reichwald, Peter Lichtner, André Scherag, et al. 2006. “Prevalence, Spectrum, and Functional Characterization of Melanocortin-4 Receptor Gene Mutations in a Representative Population-Based Sample and Obese Adults from Germany.” *The Journal of Clinical Endocrinology and Metabolism* 91 (5): 1761–69.
Hinney, Anke, Sarah Hohmann, Frank Geller, Constanze Vogel, Claudia Hess, Anne-Kathrin Wermter, Britta Brokamp, et al. 2003. “Melanocortin-4 Receptor Gene: Case-Control Study and Transmission Disequilibrium Test Confirm That Functionally Relevant Mutations Are Compatible with a Major Gene Effect for Extreme Obesity.” *The Journal of Clinical Endocrinology and Metabolism* 88 (9): 4258–67.
Hinney, Anke, Antje Körner, and Pamela Fischer-Posovszky. 2022. “The Promise of New Anti-Obesity Therapies Arising from Knowledge of Genetic Obesity Traits.” *Nature Reviews. Endocrinology*, July, 1–15.
Hinney, Anke, Anna-Lena Volckmar, and Nadja Knoll. 2013. “Melanocortin-4 Receptor in Energy Homeostasis and Obesity Pathogenesis.” *Progress in Molecular Biology and Translational Science* 114:147–91.
Howard, Matthew K., Nicholas Hoppe, Xi-Ping Huang, Christian B. Macdonald, Eshan Mehrota, Patrick Rockefeller Grimes, Adam Zahm, et al. 2024. “Molecular Basis of Proton-Sensing by G Protein-Coupled Receptors.” *bioRxiv : The Preprint Server for Biology*, April. https://doi.org/10.1101/2024.04.17.590000.
Hruby, V. J., B. C. Wilkes, M. E. Hadley, F. Al-Obeidi, T. K. Sawyer, D. J. Staples, A. E. de Vaux, O. Dym, A. M. Castrucci, and M. F. Hintz. 1987. “Alpha-Melanotropin: The Minimal Active Sequence in the Frog Skin Bioassay.” *Journal of Medicinal Chemistry* 30 (11): 2126–30.
Huang, Hui, and Ya-Xiong Tao. 2014. “A Small Molecule Agonist THIQ as a Novel Pharmacoperone for Intracellularly Retained Melanocortin-4 Receptor Mutants.” *International Journal of Biological Sciences* 10 (8): 817–24.
Huang, Hui, Wei Wang, and Ya-Xiong Tao. 2017. “Pharmacological Chaperones for the Misfolded Melanocortin-4 Receptor Associated with Human Obesity.” *Biochimica et Biophysica Acta, Molecular Basis of Disease* 1863 (10 Pt A): 2496–2507.
Hughes, J. P., S. Rees, S. B. Kalindjian, and K. L. Philpott. 2011. “Principles of Early Drug Discovery.” *British Journal of Pharmacology* 162 (6): 1239–49.
Isberg, Vignir, Chris de Graaf, Andrea Bortolato, Vadim Cherezov, Vsevolod Katritch, Fiona H. Marshall, Stefan Mordalski, et al. 2015. “Generic GPCR Residue Numbers - Aligning Topology Maps While Minding the Gaps.” *Trends in Pharmacological Sciences* 36 (1): 22–31.
Jones, Eric M., Rishi Jajoo, Daniel Cancilla, Nathan B. Lubock, Jeffrey Wang, Megan Satyadi, Rockie Chong, et al. 2019. “A Scalable, Multiplexed Assay for Decoding GPCR-Ligand Interactions with RNA Sequencing.” *Cell Systems* 8 (3): 254–60.e6.
Jones, Eric M., Nathan B. Lubock, A. J. Venkatakrishnan, Jeffrey Wang, Alex M. Tseng, Joseph M. Paggi, Naomi R. Latorraca, et al. 2020. “Structural and Functional Characterization of G Protein-Coupled Receptors with Deep Mutational Scanning.” *eLife* 9 (October). https://doi.org/10.7554/eLife.54895.
Ju, Sang Hyeon, Gyu-Bon Cho, and Jong-Woo Sohn. 2018. “Understanding Melanocortin-4 Receptor Control of Neuronal Circuits: Toward Novel Therapeutics for Obesity Syndrome.” *Pharmacological Research: The Official Journal of the Italian Pharmacological Society* 129 (March):10–19.

Kievit, Paul, Heather Halem, Daniel L. Marks, Jesse Z. Dong, Maria M. Glavas, Puspha Sinnayah, Lindsay Pranger, Michael A. Cowley, Kevin L. Grove, and Michael D. Culler. 2013. “Chronic Treatment with a Melanocortin-4 Receptor Agonist Causes Weight Loss, Reduces Insulin Resistance, and Improves Cardiovascular Function in Diet-Induced Obese Rhesus Macaques.” *Diabetes* 62 (2): 490–97.

Lafita, Aleix, Ferran Gonzalez, Mahmoud Hossam, Paul Smyth, Jacob Deasy, Ari Allyn-Feuer, Daniel Seaton, and Stephen Young. 2024. “Fine-Tuning Protein Language Models with Deep Mutational Scanning Improves Variant Effect Prediction.” *arXiv [q-bio.GN]*. arXiv. http://arxiv.org/abs/2405.06729.

Landrum, Melissa J., Jennifer M. Lee, George R. Riley, Wonhee Jang, Wendy S. Rubinstein, Deanna M. Church, and Donna R. Maglott. 2014. “ClinVar: Public Archive of Relationships among Sequence Variation and Human Phenotype.” *Nucleic Acids Research* 42 (Database issue): D980–85.

Lenth, Russell V. 2024. “Estimated Marginal Means, Aka Least-Squares Means [R Package Emmeans Version 1.10.1],” April. https://CRAN.R-project.org/package=emmeans.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79.

Li, Yong-Qi, Yogendra Shrestha, Mritunjay Pandey, Min Chen, Ahmed Kablan, Oksana Gavrilova, Stefan Offermanns, and Lee S. Weinstein. 2016. “G(q/11)α and G(s)α Mediate Distinct Physiological Responses to Central Melanocortins.” *The Journal of Clinical Investigation* 126 (1): 40–49.

Lotta, Luca A., Jacek Mokrosiński, Edson Mendes de Oliveira, Chen Li, Stephen J. Sharp, Jian ’an Luan, Bas Brouwers, et al. 2019. “Human Gain-of-Function MC4R Variants Show Signaling Bias and Protect against Obesity.” *Cell* 177 (3): 597–607.e9.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.

Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBnet.journal* 17 (1): 10–12.

Matreyek, Kenneth A., Lea M. Starita, Jason J. Stephany, Beth Martin, Melissa A. Chiasson, Vanessa E. Gray, Martin Kircher, et al. 2018. “Multiplex Assessment of Protein Variant Abundance by Massively Parallel Sequencing.” *Nature Genetics* 50 (6): 874–82.

McCarthy, Davis J., Yunshun Chen, and Gordon K. Smyth. 2012. “Differential Expression Analysis of Multifactor RNA-Seq Experiments with Respect to Biological Variation.” *Nucleic Acids Research* 40 (10): 4288–97.

McDonald, Eli Fritz, Kathryn E. Oliver, Jonathan P. Schlebach, Jens Meiler, and Lars Plate. 2024. “Benchmarking AlphaMissense Pathogenicity Predictions against Cystic Fibrosis Variants.” *PloS One* 19 (1): e0297560.

Metzger, Peter J., Aileen Zhang, Bradley A. Carlson, Hui Sun, Zhenzhong Cui, Yongqi Li, Marshal T. Jahnke, et al. 2024. “A Human Obesity-Associated MC4R Mutation with Defective Gq/11α Signaling Leads to Hyperphagia in Mice.” *The Journal of Clinical Investigation* 134 (4). https://doi.org/10.1172/JCI165418.

Notin, Pascal, Aaron W. Kollasch, Daniel Ritter, Lood van Niekerk, Steffanie Paul, Hansen Spinner, Nathan Rollins, et al. 2023. “ProteinGym: Large-Scale Benchmarks for Protein Design and Fitness Prediction.” *bioRxiv : The Preprint Server for Biology*, December. https://doi.org/10.1101/2023.12.07.570727.

Orenbuch, Rose, Aaron W. Kollasch, Hansen D. Spinner, Courtney A. Shearer, Thomas A. Hopf, Dinko Franceschi, Mafalda Dias, Jonathan Frazer, and Debora S. Marks. 2023. “Deep Generative Modeling of the Human Proteome Reveals over a Hundred Novel Genes Involved in Rare Genetic Disorders.” *medRxiv : The Preprint Server for Health Sciences*,

November. https://doi.org/10.1101/2023.11.27.23299062.
Paisdzior, Sarah, Ioanna Maria Dimitriou, Paul Curtis Schöpe, Paolo Annibale, Patrick Scheerer, Heiko Krude, Martin J. Lohse, Heike Biebermann, and Peter Kühnen. 2020. “Differential Signaling Profiles of MC4R Mutations with Three Different Ligands.” *International Journal of Molecular Sciences* 21 (4). https://doi.org/10.3390/ijms21041224.
Pettersen, Eric F., Thomas D. Goddard, Conrad C. Huang, Elaine C. Meng, Gregory S. Couch, Tristan I. Croll, John H. Morris, and Thomas E. Ferrin. 2021. “UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers.” *Protein Science: A Publication of the Protein Society* 30 (1): 70–82.
Plenge, Robert M., Edward M. Scolnick, and David Altshuler. 2013. “Validating Therapeutic Targets through Human Genetics.” *Nature Reviews. Drug Discovery* 12 (8): 581–94.
Podyma, Brandon, Hui Sun, Eric A. Wilson, Bradley Carlson, Ethan Pritikin, Oksana Gavrilova, Lee S. Weinstein, and Min Chen. 2018. “The Stimulatory G Protein Gsα Is Required in Melanocortin 4 Receptor-Expressing Cells for Normal Energy Balance, Thermogenesis, and Glucose Metabolism.” *The Journal of Biological Chemistry* 293 (28): 10993–5.
Poitout, Lydie, Valérie Brault, Carole Sackur, Sonia Bernetière, José Camara, Pascale Plas, and Pierre Roubert. 2007. “Identification of a Novel Series of Benzimidazoles as Potent and Selective Antagonists of the Human Melanocortin-4 Receptor.” *Bioorganic & Medicinal Chemistry Letters* 17 (16): 4464–70.
Quan, Jenai, Charles Langelier, Alison Kuchta, Joshua Batson, Noam Teyssier, Amy Lyden, Saharai Caldera, et al. 2019. “FLASH: A next-Generation CRISPR Diagnostic for Multiplexed Detection of Antimicrobial Resistance Sequences.” *Nucleic Acids Research* 47 (14): e83.
Robinson, Mark D., and Alicia Oshlack. 2010. “A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data.” *Genome Biology* 11 (3): R25.
Robinson, Mark D., and Gordon K. Smyth. 2007. “Moderated Statistical Tests for Assessing Differences in Tag Abundance.” *Bioinformatics* 23 (21): 2881–87.
Rodríguez Rondón, Alejandra V., Mila S. Welling, Erica L. T. van den Akker, Elisabeth F. C. van Rossum, Elles M. J. Boon, Mieke M. van Haelst, Patric J. D. Delhanty, and Jenny A. Visser. 2024. “MC4R Variants Modulate α-MSH and Setmelanotide Induced Cellular Signaling at Multiple Levels.” *The Journal of Clinical Endocrinology and Metabolism*, April. https://doi.org/10.1210/clinem/dgae210.
Rubin, Alan F., Hannah Gelman, Nathan Lucas, Sandra M. Bajjalieh, Anthony T. Papenfuss, Terence P. Speed, and Douglas M. Fowler. 2017. “A Statistical Framework for Analyzing Deep Mutational Scanning Data.” *Genome Biology* 18 (1): 150.
Sebhat, Iyassu K., William J. Martin, Zhixiong Ye, Khaled Barakat, Ralph T. Mosley, David B. R. Johnston, Raman Bakshi, et al. 2002. “Design and Pharmacology of N-[(3R)-1,2,3,4-Tetrahydroisoquinolinium- 3-Ylcarbonyl]-(1R)-1-(4-Chlorobenzyl)-2-[4-Cyclohexyl-4-(1H-1,2,4-Triazol- 1-Ylmethyl)piperidin-1-Yl]-2-Oxoethylamine (1), a Potent, Selective, Melanocortin Subtype-4 Receptor Agonist.” *Journal of Medicinal Chemistry* 45 (21): 4589–93.
Sharma, Shubh, Alastair S. Garfield, Bhavik Shah, Patrick Kleyn, Ilia Ichetovkin, Ida Hatoum Moeller, William R. Mowrey, and Lex H. T. Van der Ploeg. 2019. “Current Mechanistic and Pharmacodynamic Understanding of Melanocortin-4 Receptor Activation.” *Molecules* 24 (10). https://doi.org/10.3390/molecules24101892.
Smith, Tom, Andreas Heger, and Ian Sudbery. 2017. “UMI-Tools: Modeling Sequencing Errors in Unique Molecular Identifiers to Improve Quantification Accuracy.” *Genome Research* 27 (3): 491–99.
Sriram, Krishna, and Paul A. Insel. 2018. “G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs?” *Molecular Pharmacology* 93 (4): 251–58.

Starita, Lea M., Nadav Ahituv, Maitreya J. Dunham, Jacob O. Kitzman, Frederick P. Roth, Georg Seelig, Jay Shendure, and Douglas M. Fowler. 2017. “Variant Interpretation: Functional Assays to the Rescue.” *American Journal of Human Genetics* 101 (3): 315–25.

Stutzmann, Fanny, Karen Tan, Vincent Vatin, Christian Dina, Béatrice Jouret, Jean Tichet, Beverley Balkau, et al. 2008. “Prevalence of Melanocortin-4 Receptor Deficiency in Europeans and Their Age-Dependent Penetrance in Multigenerational Pedigrees.” *Diabetes* 57 (9): 2511–18.

Sweeney, Patrick, Luis E. Gimenez, Ciria C. Hernandez, and Roger D. Cone. 2023. “Targeting the Central Melanocortin System for the Treatment of Metabolic Disorders.” *Nature Reviews. Endocrinology* 19 (9): 507–19.

Tao, Ya-Xiong. 2010. “The Melanocortin-4 Receptor: Physiology, Pharmacology, and Pathophysiology.” *Endocrine Reviews* 31 (4): 506–43.

Vaisse, C., K. Clement, E. Durand, S. Hercberg, B. Guy-Grand, and P. Froguel. 2000. “Melanocortin-4 Receptor Mutations Are a Frequent and Heterogeneous Cause of Morbid Obesity.” *The Journal of Clinical Investigation* 106 (2): 253–62.

Wade, Kaitlin H., Brian Y. H. Lam, Audrey Melvin, Warren Pan, Laura J. Corbin, David A. Hughes, Kara Rainbow, et al. 2021. “Loss-of-Function Mutations in the Melanocortin 4 Receptor in a UK Birth Cohort.” *Nature Medicine* 27 (6): 1088–96.

Wang, Xiao-Hua, Hao-Meng Wang, Bao-Lei Zhao, Peng Yu, and Zhen-Chuan Fan. 2014. “Rescue of Defective MC4R Cell-Surface Expression and Signaling by a Novel Pharmacoperone Ipsen 17.” *Journal of Molecular Endocrinology* 53 (1): 17–29.

Weaver, James L., Wendy Wu, Paula L. Hyland, Robert Lim, Patroula Smpokou, and Michael Pacanowski. 2022. “Expanding Approved Patient Populations for Rare Disease Treatment Using In Vitro Data.” *Clinical Pharmacology and Therapeutics* 112 (1): 58–61.

Weile, Jochen, and Frederick P. Roth. 2018. “Multiplexed Assays of Variant Effects Contribute to a Growing Genotype-Phenotype Atlas.” *Human Genetics* 137 (9): 665–78.

Weng, Chenchun, Andre J. Faure, Albert Escobedo, and Ben Lehner. 2024. “The Energetic and Allosteric Landscape for KRAS Inhibition.” *Nature* 626 (7999): 643–52.

Yeo, Giles S. H., Emma J. Lank, I. Sadaf Farooqi, Julia Keogh, Benjamin G. Challis, and Stephen O’Rahilly. 2003. “Mutations in the Human Melanocortin-4 Receptor Gene Associated with Severe Familial Obesity Disrupts Receptor Function through Multiple Molecular Mechanisms.” *Human Molecular Genetics* 12 (5): 561–74.

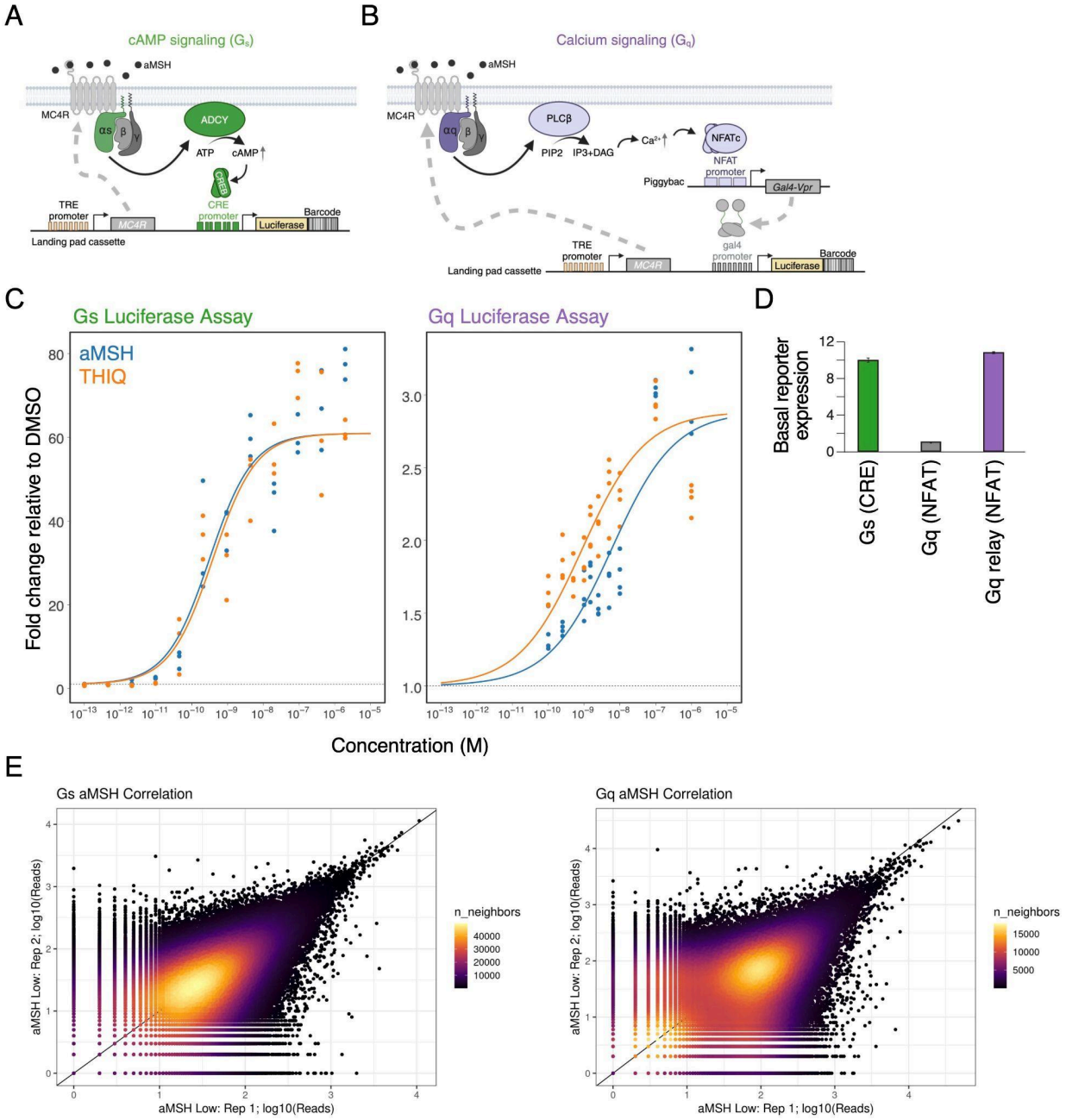
Yu, Jing, Luis E. Gimenez, Ciria C. Hernandez, Yiran Wu, Ariel H. Wein, Gye Won Han, Kyle McClary, et al. 2020. “Determination of the Melanocortin-4 Receptor Structure Identifies Ca2+ as a Cofactor for Ligand Binding.” *Science* 368 (6489): 428–33.

Zhang, Huibing, Li-Nan Chen, Dehua Yang, Chunyou Mao, Qingya Shen, Wenbo Feng, Dan-Dan Shen, et al. 2021. “Structural Insights into Ligand Recognition and Activation of the Melanocortin-4 Receptor.” *Cell Research* 31 (11): 1163–75.

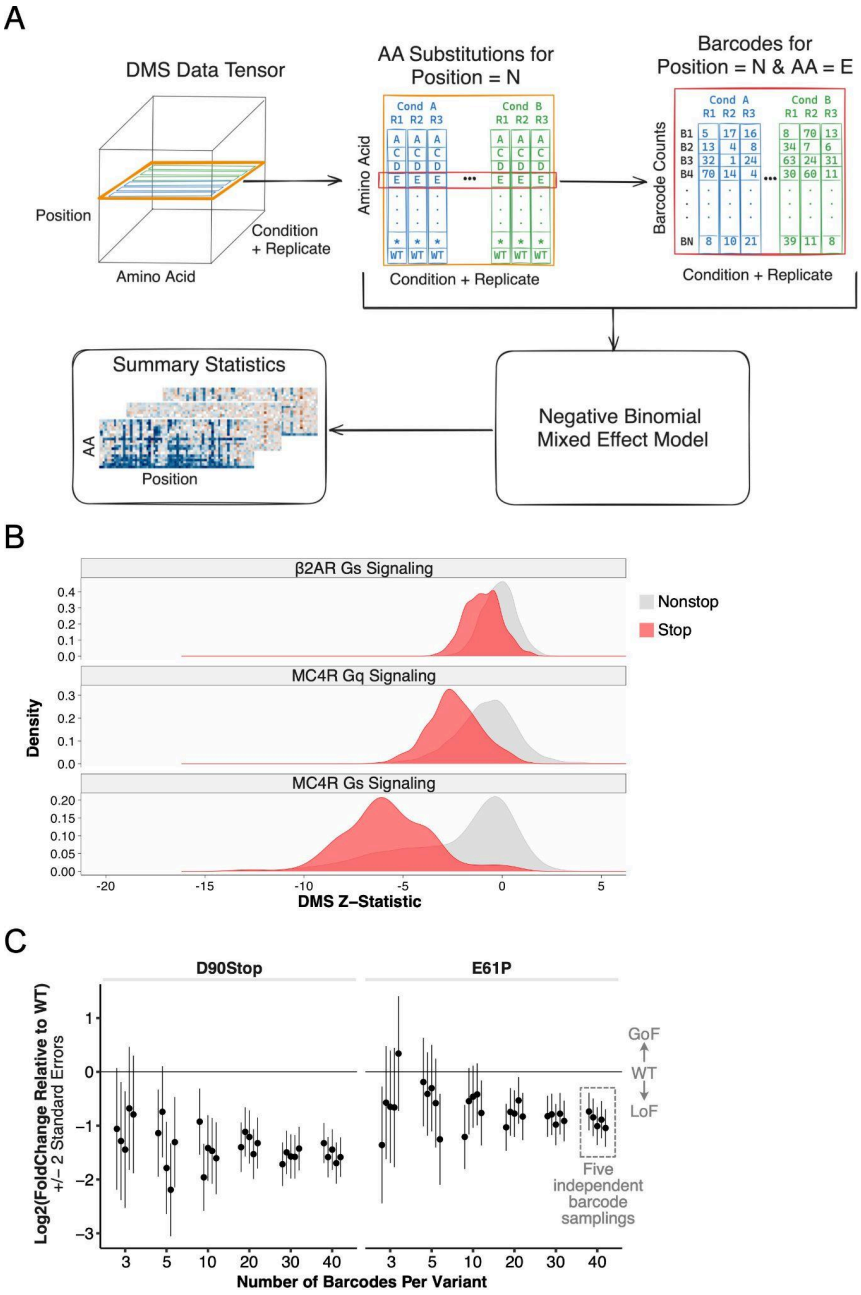
Zhou, Qingtong, Dehua Yang, Meng Wu, Yu Guo, Wanjing Guo, Li Zhong, Xiaoqing Cai, et al. 2019. “Common Activation Mechanism of Class A GPCRs.” *eLife* 8 (December). https://doi.org/10.7554/eLife.50279.

Supplemental Information

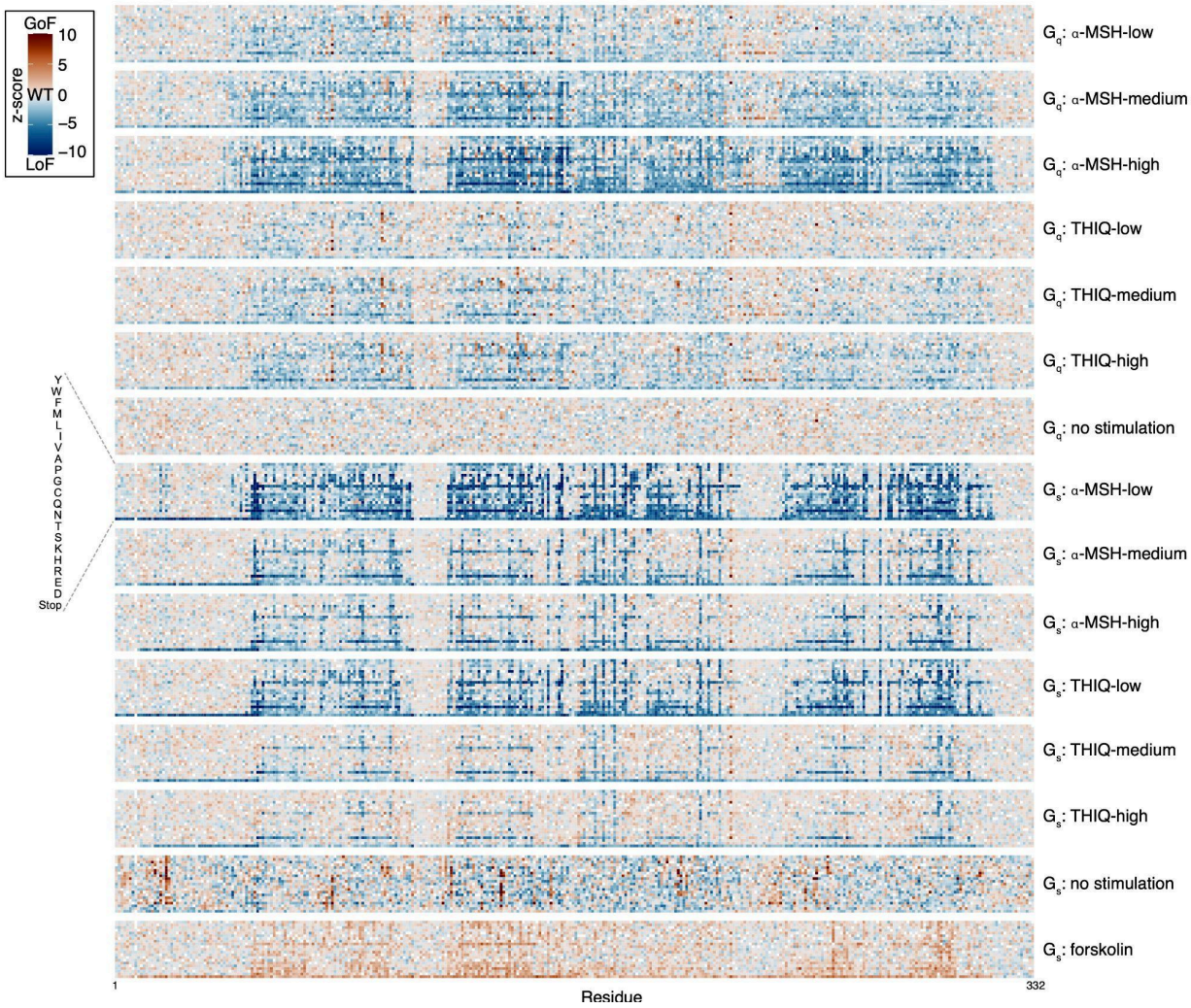
Supplemental Figures



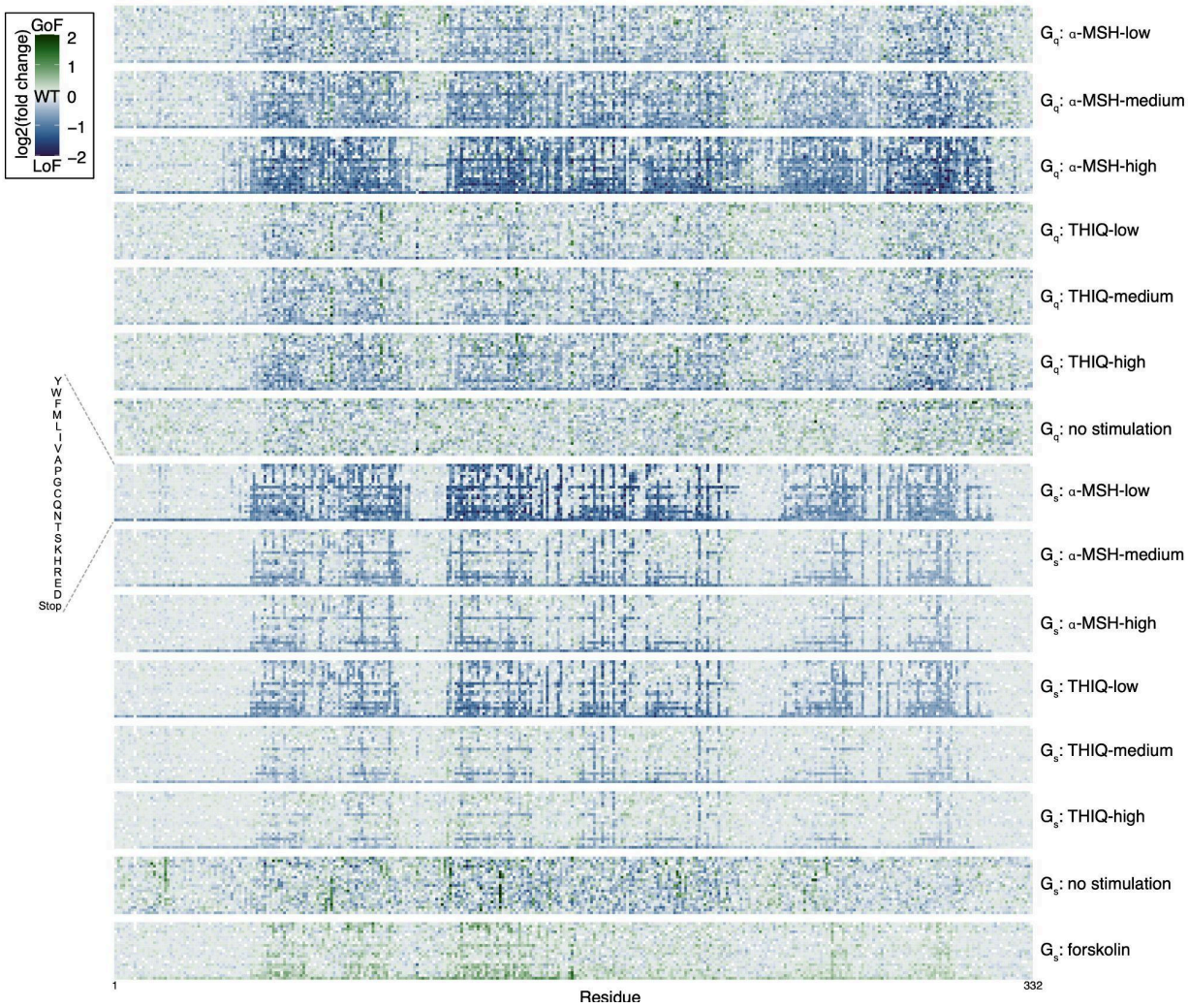
Supplementary Figure 1. Methods development detail. A, B. Detailed cartoons of reporters for cAMP (Gs) signaling (**A**) and calcium (Gq) signaling (**B**). **C.** Luciferase assays used to determine agonist (α-MSH and THIQ) dose ranges for the DMS experiments. For each reporter (Gs, Gq), we fit a 3-parameter log-logistic with the bottom of the curve fixed to the DMSO mean. We shared the slope and Emax parameter between α-MSH and THIQ, and let the EC50 parameter vary between the drugs. **D.** Basal reporter gene expression for the Gs (CRE) reporter, the Gq (NFAT) reporter, and the Gq relay (NFAT) reporter. **E.** Correlation of sequence read counts of each barcode for two experimental replicates each of the Gs (**left**) and Gq (**right**) DMS assays.



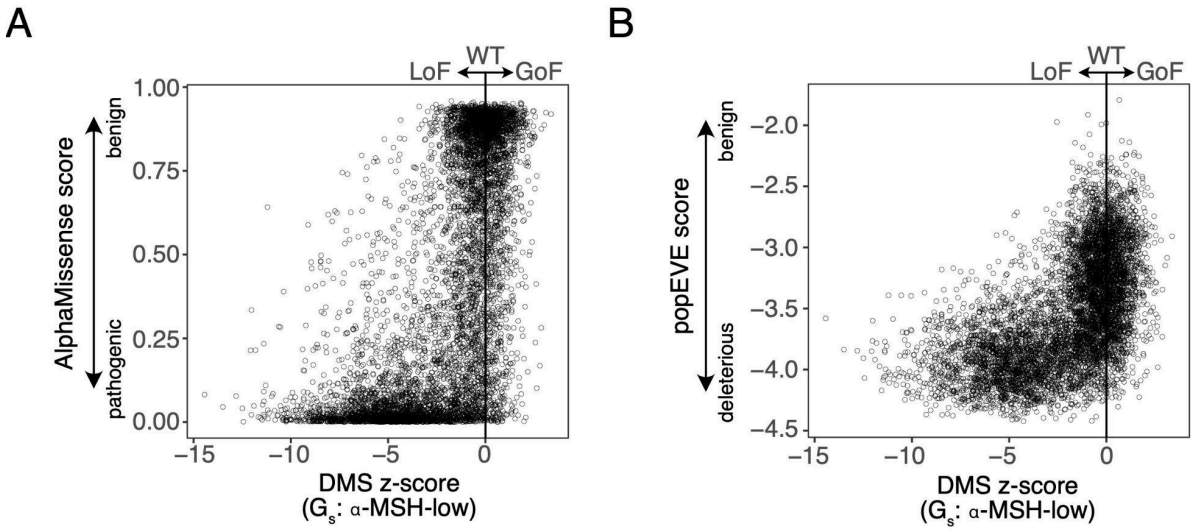
Supplementary Figure 2. Analysis method and barcoding detail. A. Schematic for analysis method. DMS data can be represented as a tensor, with dimensions for Position in the protein, Amino Acid substitution, and Condition + Replicate. Within the Condition + Replicate dimension, there are multiple drugs (i.e., α-MSH and THIQ) and multiple signaling pathways (i.e., Gs and Gq). Furthermore, each amino acid substitution has multiple barcodes associated with it. Our negative binomial mixed-effect model takes all of this structure into account to produce summary statistics that are the basis of our comparisons. **B.** Distribution of Z-statistics for stop codons versus all other variant effects from published β2AR (Jones et al. 2020) and MC4R (this study) DMS assays assessing Gs or Gq signaling activity. **C.** Downsampling simulation from the MC4R data showing the effect of number of barcodes per variant on the estimation of log2 fold change relative to wild type. Here we show two representative positions in the α-MSH low condition, randomly downsampled a fixed number of barcodes associated with that variant. We repeated that process for 5 independent samplings, and ran the resulting data through our model resulting in parameter estimates. Increasing the number of barcodes per variant decreases the magnitude of the error bars (+/- 2 standard errors).



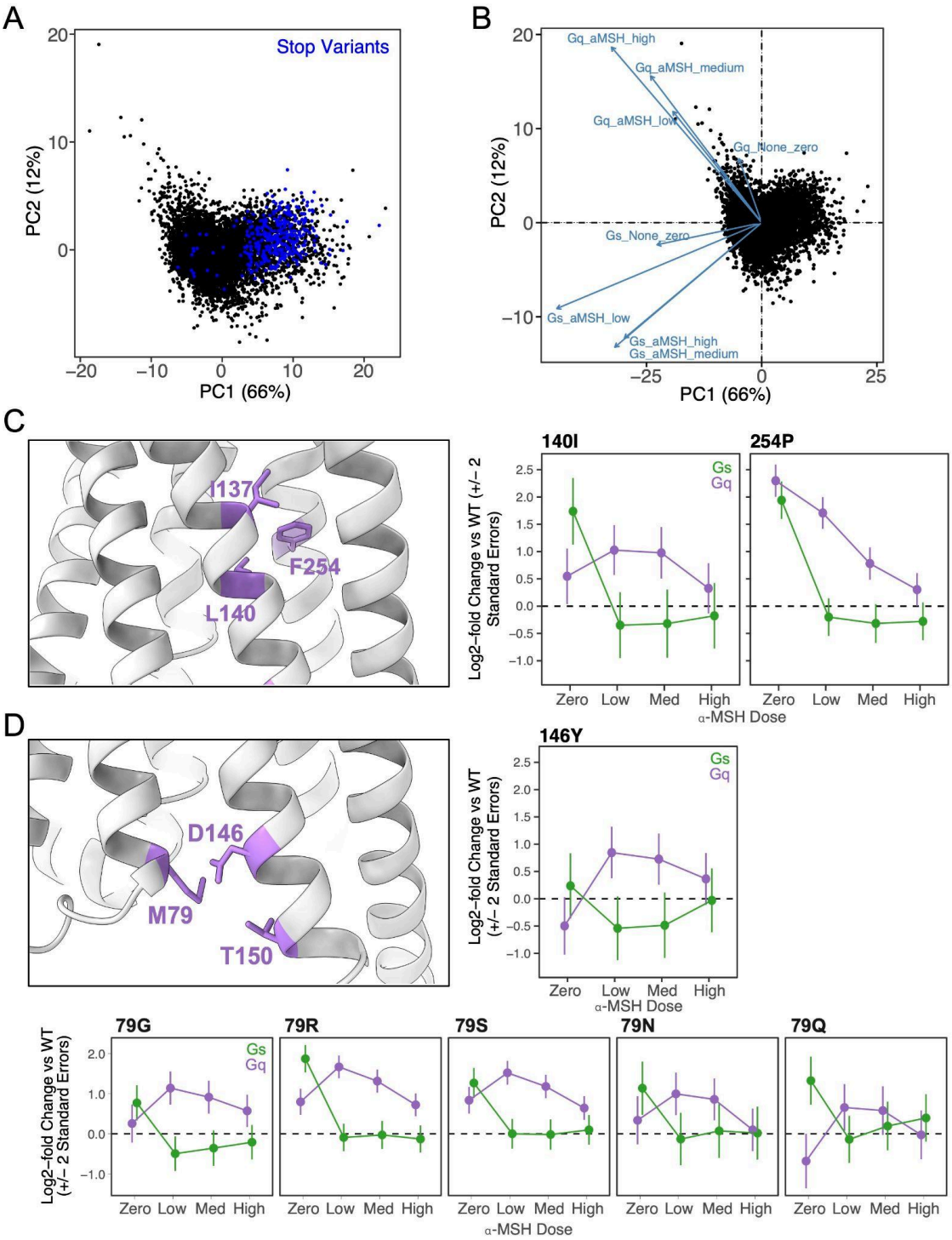
Supplementary Figure 3. Results (z-scores) of MC4R Deep Mutational Scans across all agonist stimulation conditions. Shown are z-scores indicating the effect of each variant (y-axis) at each position (x-axis) relative to wild-type. GoF: gain-of-function; LoF: loss-of-function, WT: wild-type.



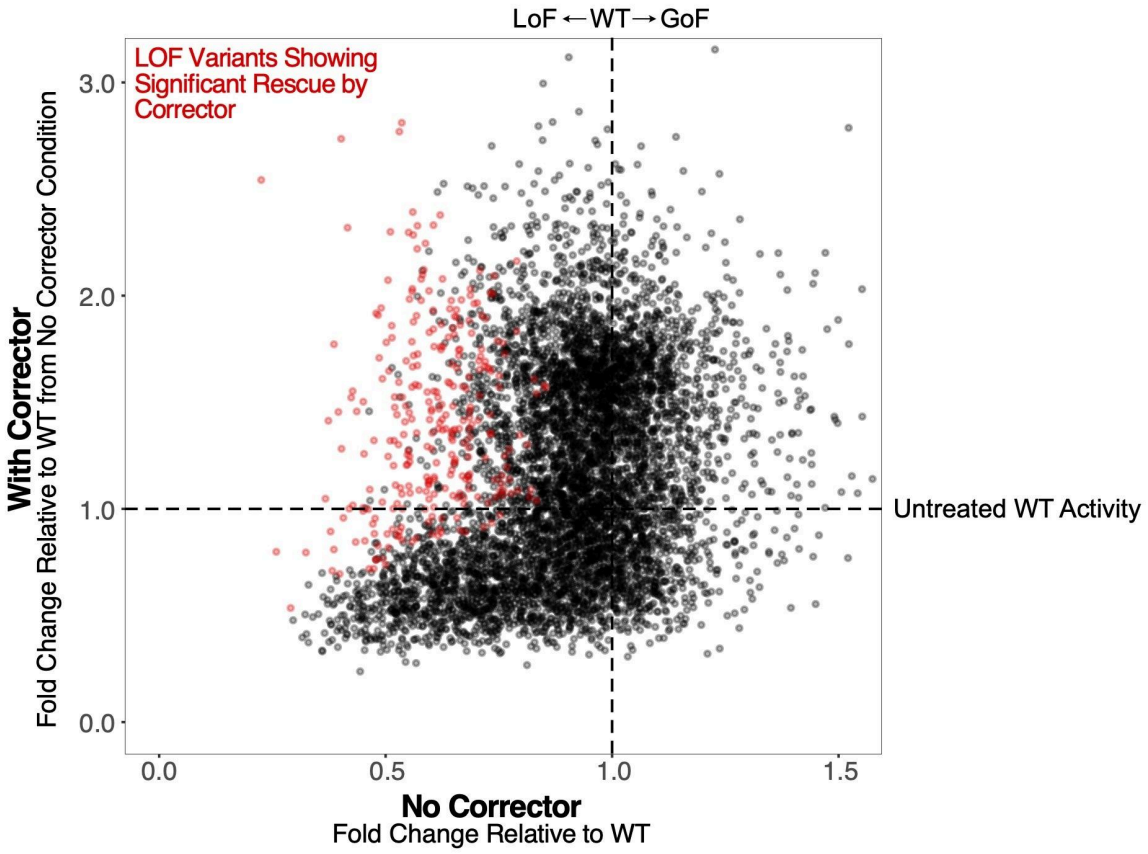
Supplementary Figure 4. Results (log_2[fold change]) of MC4R Deep Mutational Scans across all agonist stimulation conditions. Shown are the log_2(fold change) values of each variant (y-axis) at each position (x-axis) relative to wild-type. GoF: gain-of-function; LoF: loss-of-function, WT: wild-type.



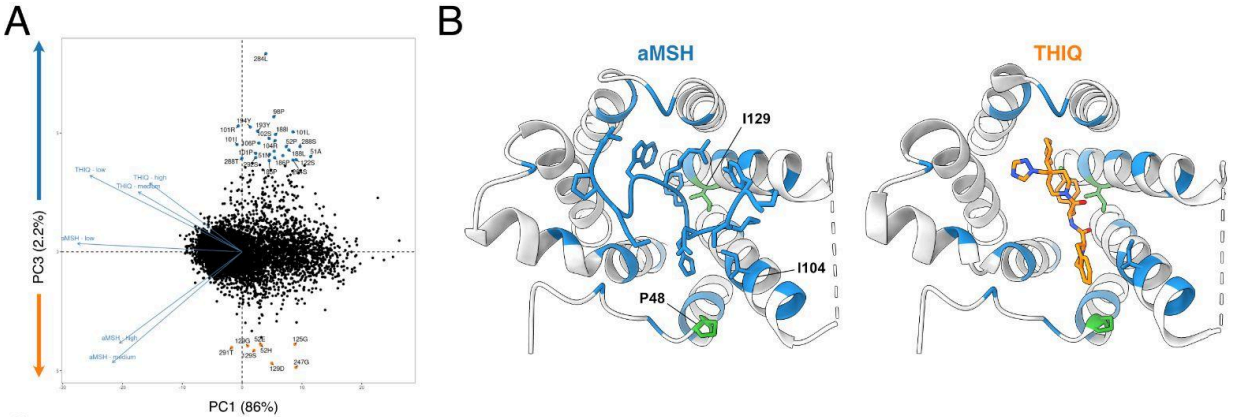
Supplementary Figure 5. MC4R Deep Mutational Scans are consistent with variant effect predictions. A., B. Shown are z-scores for each MC4R variant (x-axis) from the Gs-α-MSH-low (forskolin normalized) DMS condition relative to the AlphaMissense (**A**) or popEve (**B**) score of the variant. GoF: gain-of-function; LoF: loss-of-function, WT: wild-type.



Supplementary Figure 6. Variants with differential effects on Gs and Gq signaling. A. PCA of z-scores from all Gs and Gq assays using no, low, medium, and high α-MSH. Each point represents a variant, with stop codon variants colored in blue and supporting that PC1 is driven by variants that are LoF in both Gs and Gq assays versus those that retain activity in one or both assays. **B.** Same PCA as **A** but with arrows showing how each of the incorporated datasets is contributing to the components. This supports that PC2 is driven by differences in Gs and Gq signaling. **C,D.** Additional examples of variants that have differential effects on Gs and Gq signal, as in **Fig. 3C,D**. **Left:** Closeup of MC4R structure showing positions of interest in purple. **Right/Bottom:** MC4R signaling activity (log2[fold change of variant activity relative to wild-type]) for selected variants across α-MSH doses. Error bars are +/- 2 standard errors. Med: medium; WT: wild-type



Supplementary Figure 7. Identification of variants that respond to corrector treatment. Scatter plot showing activity of each MC4R variant in Gs signaling assay upon stimulation with 1 μM α-MSH (>EC99) in the absence of corrector (x-axis) versus upon treatment with 1 μM Ipsen 17 followed by stimulation with 1 μM α-MSH (>EC99) (y-axis). Each point represents a variant, with those in red classified as both 1) loss-of-function in the absence of corrector; and 2) functionally rescued in the Ipsen 17 treatment condition (see **Methods** for details).



Supplementary Figure 8. Variants that differentially impact activation by peptide versus small molecule agonists. A. PCA of z-scores from all Gs signaling assays with no, low, medium, and high concentrations of α-MSH or THIQ. Each point represents a variant colored by whether it uniquely abrogates signaling in the α-MSH (blue) or THIQ (orange) condition(s). **B.** Top-down ribbon view of MC4R with bound α-MSH (left, blue; PDB: 7f53) or THIQ (right, orange; PDB: 7f58). MC4R residues are colored by whether variants at a given position disrupt activation uniquely by α-MSH (blue), THIQ (orange), both (green), or neither (gray). Residues 48, 104, and 129 are shown in stick form.

Supplemental Tables

The following supplementary tables are provided as a separate Excel file:

Supplementary Table 1. **Experimental conditions across all DMS assays**

Supplementary Table 2. Functional effects of >200 human MC4R variants

Supplementary Table 3. Oligo sequences used in this study