Structural Biology and Molecular Biophysics

Divergent Pairwise Epistasis in the Context of Unstable Membrane Protein Variants

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint version 1
December 1, 2023 (this version)

Sent for peer review
September 14, 2023

Posted to preprint server
August 27, 2023

Laura M. Chamness, Charles P. Kuntz, Andrew G. McKee, Wesley D. Penn, Christopher M. Hemmerich, Douglas B. Rusch, Hope Woods, Dyotima, Jens Meiler, Jonathan P. Schlebach

Department of Chemistry, Indiana University, Bloomington, Indiana, USA • Department of Chemistry, Purdue University, West Lafayette, Indiana, USA • Center for Genomics and Bioinformatics, Indiana University, Bloomington, Indiana, USA • Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA • Chemical and Physical Biology Program, Vanderbilt University, Nashville, Tennessee, USA • Institute for Drug Development, Leipzig University, Leipzig, SAC, Germany

Abstract

Many eukaryotic membrane proteins are prone to misfolding, which compromises their functional expression at the plasma membrane. This is particularly true for mammalian gonadotropin-releasing hormone receptors (GnRHRs), which are G protein-coupled receptors involved in reproductive steroidogenesis. We recently demonstrated that evolutionary modifications within mammalian GnRHRs appear to have coincided with adaptive changes in cotranslational folding efficiency. Though changes in protein stability are known to shape evolutionary interactions, it is unclear how the energetic drivers of cotranslational folding in the membrane may modify epistatic interactions. We therefore surveyed the pairwise epistatic interactions that modify the expression of two destabilized GnRHR variants bearing mutations that selectively compromise either its membrane topology (V276T) or its native tertiary structure (W107A). Using deep mutational scanning (DMS), we evaluated how the effects of these mutations on the expression of the mature form of the protein at the plasma membrane are modified by hundreds of secondary mutations. A focused analysis of 251 mutants with high-quality measurements in three genetic backgrounds reveals that V276T and W107A form distinct epistatic interactions that depend on both the degree to which they destabilize the protein and the mechanism of their destabilization. An unsupervised learning analysis shows that V276T forms predominantly negative epistatic interactions that are most pronounced among destabilizing mutations within soluble loop regions. In contrast, W107A forms interactions with mutations in both loops and transmembrane domains that skew positive as a result of the diminishing impact of the destabilizing mutations in the context of an already unstable variant. These findings provide general insights into how pairwise epistasis is remodeled by conformational defects in membrane proteins and, more generally, in unstable proteins.

eLife assessment

The **important** study describes exhaustive deep mutational scanning (DMS) of the gonadotropin-releasing hormone wild-type receptor and for two single point mutations reported to impact its folding and structure, monitoring how plasma membrane expression levels are affected by mutations. This **important** work is pioneering in exploring the interaction between mutations (epistasis) in a membrane protein, with a potential for explaining membrane protein evolution and genetic diseases. The evidence provided for some mutations is **convincing**, but it remains **incomplete** and harder to interpret for others without further validation of folding and stability properties of the mutants.

Introduction

Mutational effects on protein stability have important consequences for evolution. Destabilized proteins misfold more often, which can attenuate their ability to function in the cell (4, 17). Though natural selection typically maximizes fitness by incorporating mutations that produce new or improved functions, this optimization process is often hindered by the destabilizing effects of most random mutations (4, 39). The cumulative energetic effects of these mutations on protein stability are generally capable of creating non-additive, context-dependent epistatic interactions (10, 26, 36). Nevertheless, there are many different aspects of protein synthesis, folding, and assembly that are governed by distinct energetic constraints. While epistatic interactions arising from changes in protein stability have been previously characterized in soluble proteins (Gong et al. 2013; Olson et al. 2014; Faber et al. 2019; Nedrud et al. 2021), the impact of epistasis on membrane protein folding and stability has yet to be explored. Although many of the mechanistic underpinnings of pairwise epistasis in soluble proteins are likely to be generalizable to all proteins, membrane proteins undergo additional cotranslational folding reactions that are governed by mechanistically distinct kinetic and energetic constraints (17, 31). Based on these considerations, we suspect that mutations in membrane proteins could potentially modify cotranslational processes in a manner that generates distinct epistatic interactions that bias their evolutionary pathways in unique ways.

Unlike water-soluble proteins, membrane proteins must be cotranslationally inserted into lipid bilayers (Stage I folding) in order to fold into their native structure (Stage II folding) (30). In eukaryotes, Stage I folding is primarily facilitated by the Sec61 translocon complex in concert with its various other accessory subunits (17, 33).

Ribosomes translating membrane proteins are targeted to this translocon complex at the endoplasmic reticulum (ER) membrane, where the nascent transmembrane domains (TMDs) partition into the lipid bilayer and establish their native topology relative to the membrane (Stage I) (28, 30, 41). After translation is complete, membrane proteins fold and assemble (Stage II) in a manner that is coupled to their passage through the secretory pathway to the plasma membrane and/ or other destination organelles (17, 30, 44). Though they are governed by distinct physicochemical constraints, failures in either Stage I or Stage II folding are capable of increasing the proportion of nascent membrane proteins that are retained in the ER and prematurely degraded (17, 44). We therefore expect that epistatic interactions can arise from the cumulative energetic effects of mutations on either of these processes. Given that three-dimensional protein structures are stabilized by similar energetic principles in water and lipid bilayers, mutations that modify the fidelity of Stage II folding energetics should potentially generate long-range epistatic interactions that are comparable to those observed in soluble

proteins (17). By comparison, the topological transitions that happen during the early stages of membrane protein synthesis may be coupled to neighboring loops and helices in a manner that can play a decisive role in the formation of the native topology (13, 25, 43). Based on these considerations, we reason that the effects of mutations on each process could create distinct epistatic interactions.

To gain insights into the mechanistic basis of epistatic interactions in membrane proteins, we surveyed the effects of single and double mutants on the plasma membrane expression (PME) of the gonadotropin-releasing hormone receptor (GnRHR), a G-protein coupled receptor (GPCR) involved in reproductive steroidogenesis across many species (14, 15). We have previously demonstrated that various evolutionary sequence modifications within mammalian GnRHRs have tuned its fitness by compromising the fidelity of Stage I folding (6). Nevertheless, it remains unclear how the inefficient cotranslational folding of GnRHR reshapes its tolerance of secondary mutations. Here, we utilize deep mutational scanning (DMS) to carry out a focused analysis of 251 mutations in the background of two mouse GnRHR (mGnRHR) variants that selectively compromise either Stage I or Stage II folding. Our results show that mutations which generate Stage I and Stage II folding defects form distinct epistatic interactions throughout this receptor. An unsupervised learning analysis reveals how these interactions depend on both changes in stability and the topological context of the mutation. Together, these findings suggest that the distinct biosynthetic mechanisms of integral membrane proteins may differentially shape their fitness landscapes. The general implications of our findings for the role of protein folding and stability in protein evolution are discussed.

Results

Mutational Library Design

To compare how perturbations of Stage I and Stage II folding shape epistatic interactions, we first identified and characterized two individual mutations that are likely to selectively disrupt either the cotranslational membrane integration (V276T) or the native tertiary structure (W107A) of GnRHR. We previously showed that T277 in TMD6 of human GnRHR (hGnRHR) contributes to the inefficient translocon-mediated membrane integration of TMD6, which decreases the receptor's PME (6). A recent crystallographic structure of hGnRHR confirms that the polar side chain of T277, which corresponds to V276 in the mouse receptor (89.3% sequence identity), is exposed to the lipid bilayer and does not appear to make interhelical contacts that stabilize the native structure (**Fig. 1 A & B**) (47). Thus, mutations at this position modify the hydrophobicity of TMD6 without perturbing the native tertiary contacts that stabilize the native fold. Indeed, replacing the native valine residue in TMD6 of the mouse receptor with the threonine residue found in hGnRHR reduces the efficiency of its translocon-mediated membrane integration *in vitro* (Fig. S1, Table S1). Furthermore, this substitution decreases the PME of the mouse receptor by 65 ± 4% relative to that of wild-type mGnRHR (WT, **Fig. 2**). Consistent with previous findings, these results confirm that the inefficient translocation of this helix promotes cotranslational GnRHR misfolding (6). In contrast, substitutions at W107 within extracellular loop 1 (ECL1) preserve the hydrophobicity of TMDs and their topological energetics while disrupting a conserved hydrogen bond network that is found in a wide array of GPCR structures (**Fig. 1C**) (16). Mutating this conserved tryptophan to alanine (W107A) reduces the PME of mGnRHR by 88 ± 4% relative to WT (**Fig. 2**), which suggests that disrupting this hydrogen bond network destabilizes mGnRHR in a manner that promotes its cellular misfolding. Comparing how these mutations modify the proteostatic effects of secondary mutations will therefore reveal how changes in the fidelity of Stage I (V276T) and Stage II (W107A) folding differentially impact mutational epistasis.

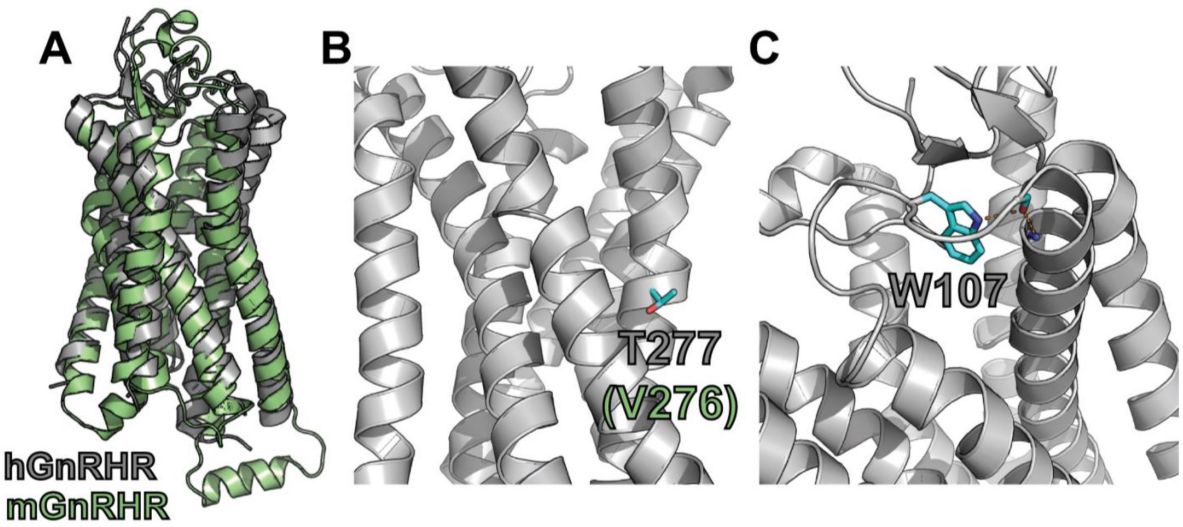


Figure 1.

Structural context of GnRHR background mutations.

A homology model of *M. musculus* GnRHR (mGnRHR) and a crystal structure of *H. sapiens* GnRHR (hGnRHR) provide structural context for two mutants which disrupt folding. A) The homology model of mGnRHR (green) structurally aligns with the crystal structure of hGnRHR (gray). B) A side view of hGnRHR shows that T277, which is analogous to V276 in mGnRHR, is located in transmembrane domain 6 and is exposed to the lipid bilayer. C) A top view of hGnRHR shows that W107 in extracellular loop 1 forms a hydrogen bond network with two other residues.

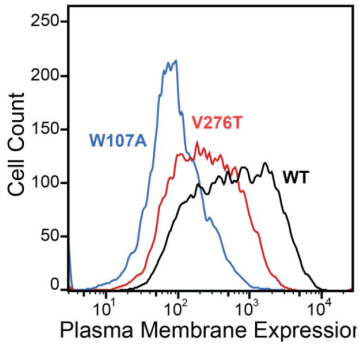


Figure 2.

Plasma Membrane Expression of mGnRHR variants.

Surface immunostaining and flow cytometry were used to compare the plasma membrane expression of select mGnRHR variants expressed in HEK293T cells. A histogram from one representative biological replicate depicts the distribution of plasma membrane expression intensities among cells expressing WT (black), V276T (red), or W107A (blue) mGnRHR.

To survey the epistatic interactions formed by these mutations, we generated a series of genetic libraries consisting of 1,615 missense variants in the background of V276T, W107A, and WT mGnRHR. To ensure adequate dynamic range in the downstream assay, we created these variants in the cDNA of mGnRHR, which exhibits intermediate, tunable expression (6). Briefly, we used a structural homology model of mGnRHR to select 85 residues distributed across the loops and helices of mGnRHR. We included all 19 amino acid substitutions at the 3 most solvent-accessible and the 3 most buried residues in each TMD (Fig. S2 and Table S2). We also included mutations encoding all 19 amino acid substitutions at positions that are evenly distributed across each soluble domain. We then generated an array of mutagenic oligonucleotides encoding each of these mutations (Table S3), which was used to introduce these mutations in the context of the V276T, W107A, and WT mGnRHR cDNAs by nicking mutagenesis (21). We generated each of these libraries in the context of plasmids containing a randomized ten-base region within the backbone. Unique 10-base plasmid identifiers that could be matched to specific mGnRHR variants of interests were used to score variants in the downstream assay using Illumina sequencing. Using whole plasmid PacBio sequencing of each library, we matched 1,383 of the 1,615 possible variants to one or more ten-base unique molecular identifier (UMI) in at least one of the three libraries, 320 of which were found in all three genetic backgrounds (see *Methods* for details).

Plasma Membrane Expression of GnRHR Mutant Libraries

To identify secondary mutations that form epistatic interactions with V276T and W107A, we measured the relative PME of single and double mutants within each genetic library by DMS as previously described (27). Briefly, we first generated a series of recombinant, pooled HEK293T cell lines in which the individual cells that each express one of the single or double mGnRHR mutants from one of the three plasmid libraries described above. A comparison of the mGnRHR surface immunostaining profiles of these cellular populations that recombinantly express this collection of variants in context of the V276T, W107A, or an otherwise WT genetic background reveals striking context-dependent differences in their proteostatic effects. Recombinant cells expressing these variants in the WT background exhibit a bimodal distribution of mGnRHR surface immunostaining, where 59% of cells exhibit immunostaining that is comparable to WT and 41% of cells exhibit reduced expression relative to WT (**Fig. 3A**). Although cells expressing this same collection of variants in the V276T background also exhibit bimodal surface immunostaining, 48% of the population exhibit diminished expression relative to the V276T variant (**Fig. 3B**). This observation suggests many of these variants synergistically decrease mGnRHR expression in combination with the V276T mutation. Unlike the WT and V276T libraries, cells expressing this same collection of variants in combination with the W107A mutation exhibit a continuous range of surface immunostaining intensities that are only slightly lower, on average, relative to cells expressing the W107A single mutant (**Fig. 3C**). The lack of dispersion likely reflects a compression of the distribution that arises as the folding efficiency of these variants approaches zero (see *Discussion*). Cells expressing this library of W107A double mutants do, however, still contain cellular subpopulations with surface immunostaining intensities that are well above or below that of the W107A single mutant, which suggests that this fluorescence signal is sensitive enough to detect subtle differences in the PME of these variants (**Fig. 3C**). Nevertheless, these results suggest that none of the secondary mutations can fully compensate for the proteostatic effects of the W107A mutation.

To compare the PME of these variants in each genetic background, we separated the mixed cell populations into quartiles based on the relative surface immunostaining of their expressed mGnRHR variants using fluorescence-activated cell sorting (FACS), then extracted the genomic DNA from each cellular fraction. We then used Illumina sequencing to track the relative abundance of the recombined UMIs and their associated variants within each fraction, then used these measurements to estimate the corresponding PME of each variant. A series of filters relating to the quality of the reads, the sampling of each mutation, and the similarity of variant scores across replicates were used to remove poorly sampled variants with unreliable scores (see *Methods*). Briefly, to restrict our analysis to variants that can be reliably scored, we removed

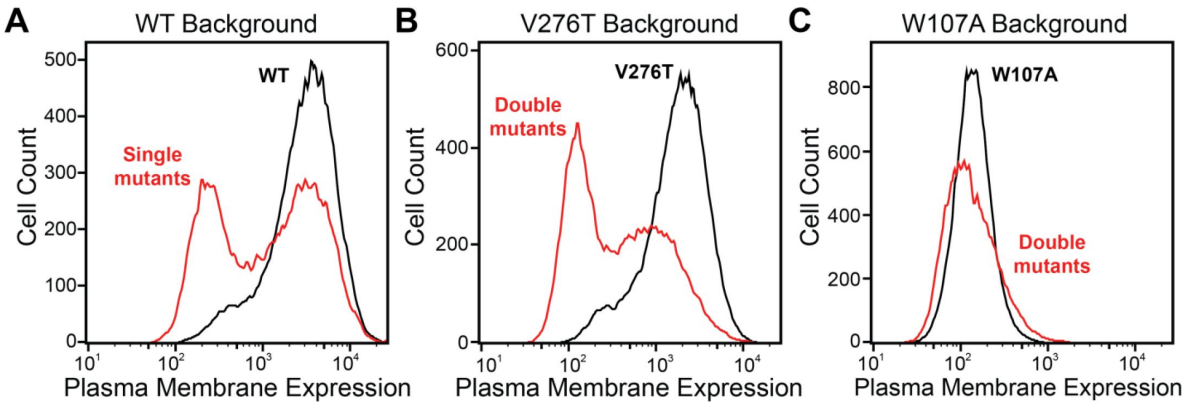


Figure 3.

Plasma membrane expression of mGnRHR cellular libraries.

mGnRHR mutant libraries (red), as well as their respective controls (WT or V276T/W107A single mutants, black) were expressed in mammalian cells, and the amount of mGnRHR expressed at the plasma membrane was measured during fluorescence-activated cell sorting. Histograms show the distributions of plasma membrane expression for mGnRHR mutants in the (A) WT library, (B) V276T library, and (C) W107A library.

variants that did not have at least 50 counts across the four bins in each replicate. Additionally, we compared the percentile rank of each variant across replicates and discarded variant scores for which the difference in percentile rank across the two replicates was greater than 25%. The relative PME of the remaining variants was averaged across two biological replicates. Overall, we measured the relative PME of 1,004 missense variants in the WT background, 338 missense variants in the V276T background, and 1,179 missense variants in the W107A background (Table S4). Out of the 320 variants identified in all three plasmid libraries by PacBio sequencing, 251 mutations passed these quality filters in the Illumina sequencing of the cellular isolates (Table S 4). In the following, we will focus on the comparison of the effects of these 251 mutations in each background in order to determine how their proteostatic effects are modified by secondary mutations.

To facilitate the comparison of estimated immunostaining intensities for each variant across the three genetic backgrounds, we normalized the raw immunostaining intensity values for each variant relative to that of the measured intensity values for the corresponding reference sequence (V276T, W107A, or WT mGnRHR). Relative intensity values of 1.0 correspond to no effect, whereas values over 1.0 correspond to mutations that enhance surface expression in that genetic background, and vice versa. A histogram of the relative intensity values for the 251 variants measured in all three backgrounds generally recapitulates the trends in the cellular immunostaining histograms (**Fig. 3** & **4A**). Most of these mutations decrease the PME of mGnRHR in all three backgrounds, which reflects the limited mutational tolerance of membrane proteins (**Fig. 4A**) (38). Notably, only a modest fraction of mutations measurably enhance the surface expression of WT mGnRHR (51 mutations, 20%) and V276T mGnRHR (21 mutations, 8% **Fig. 4A**). Mutations in the V276T background tend to decrease surface expression more than they do in the WT background, and this trend persists across most domains of the protein (**Fig. 4 A & B**). In contrast, a larger proportion of these mutations (100 mutants, 40%) enhance the surface expression of W107A (**Fig. 4A**), though these increases in immunostaining are relatively modest (**Fig. 3C**). Furthermore, the mutations tend to be better tolerated with respect to surface expression in combination with W107A relative to their effects on expression in the WT background (**Fig. 4C**).

Projecting the average relative intensities for the mutations at each residue onto the structure reveals that most positions exhibit similar trends in the mutational tolerance in each background, with loop residues being generally more permissive than those within TMDs (**Fig. 4 D – F**). Nevertheless, there are subdomains where the quantitative mutagenic effects deviate in a context-dependent manner. Variants bearing mutations within the C-terminal region (ICL3-TMD6-ECL3-TMD7) fare consistently worse in the V276T background relative to WT (**Fig. 4 B & E**). Given that V276T perturbs the cotranslational membrane integration of TMD6 (Fig. S1, Table S1), this directional bias potentially suggests that the apparent interactions between these mutations manifest during the late stages of cotranslational folding. In contrast, mutations that are better tolerated in the context of W107A mGnRHR are located throughout the structure but are particularly abundant among residues in the middle of the primary structure that form TMD4, ICL2, and ECL2 (**Fig. 4 C & F**). Together, these observations show how the proteostatic effects of mGnRHR mutations are modified by secondary mutations that differentially affect the fidelities of Stage I and Stage II folding.

Distinct Pairwise Epistasis in V276T and W107A

To compare epistatic trends in these libraries, we calculated the magnitude of the epistatic interactions these 251 mutations form with V276T and W107A by comparing their relative effects on PME of the WT, V276T, and W107A variants using a previously described relative epistasis model (product model, Olson et al. 2014). Briefly, we defined fitness as the ratio of the single/double mutant PME to that of WT mGnRHR. Epistasis scores (ε) were then determined by taking the log of the double mutant fitness value divided by the difference between the single mutant fitness values (see *Methods*). Positive ε values denote double mutants that have greater PME than

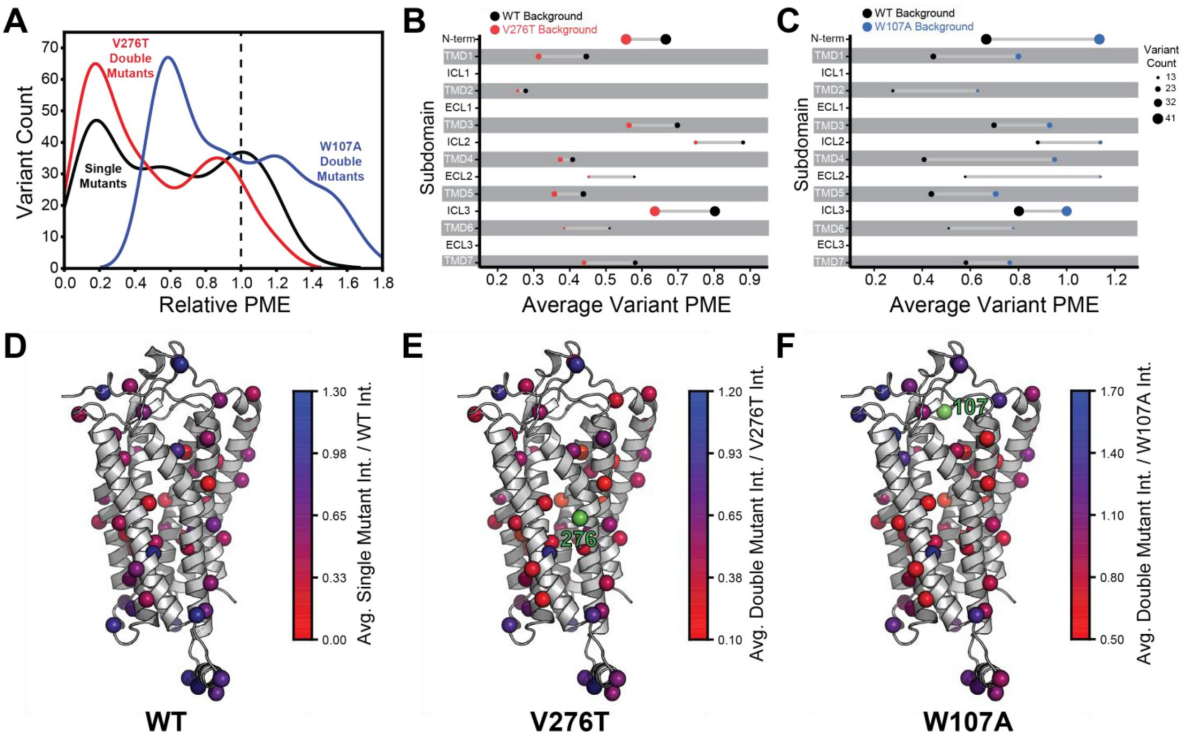


Figure 4.

Relative plasma membrane expression of mGnRHR mutants.

A) A histogram shows the distribution of plasma membrane expression measurements relative to their respective controls (WT or V276T/W107A single mutants) for the WT (gray), V276T (red) and W107A (blue) mutational libraries. A dashed line at the normalized plasma membrane expression value of 1.0, representing no mutational effect relative to the control, is displayed for reference. B & C) The average relative intensities across each domain are compared between the V276T (B, red) and W107A (C, blue) double mutant libraries and the single mutant library (gray). The size of each dot represents the number of variants measured, and significant *p*-values are also displayed. D – F) Average relative intensities for the (D) WT library, (E) V276T library, and (F) W107A library were projected onto a homology model of mGnRHR. Residues 276 and 107 are displayed in green for reference. Individual variant scores can be found in Table S4.

would be expected based on the effects of single mutants. Negative ε values denote double mutants that have lower PME than would be expected based on the effects of single mutants. Pairs of mutations with ε values near zero have additive effects on PME. Though epistatic interactions are typically rare in the context of stable proteins (36), epistasis scores among this set of 251 random mutations generally form positive epistatic interactions with W107A and negative epistatic interactions with V276T (**Fig. 5**). Interestingly, the difference in the epistasis scores for the interactions these variants form with W107A and V276T was at least 1.0 for 98 of the 251 variants (Table S4). Notably, these 98 variants are enriched with TMD variants (65% TMD) relative to the overall set of 251 variants (45% TMD). In contrast to the sparse epistasis that is generally observed between mutations within soluble proteins, these findings suggest a relatively large proportion of random mutations form epistatic interactions in the context of unstable mGnRHR variants in a manner that appears to depend on the specific folding defect (V276T vs. W107A) and topological context.

Molecular Basis for the Observed Epistatic Interactions

To identify general trends in the observed epistatic interactions, we utilized unsupervised learning to elucidate patterns among the relative PME values of variants across each genetic background. We first utilized uniform manifold approximation projection (UMAP) (20) to identify mutations that have similar expression profiles across these conditions. A projection of the variants onto a resulting two-dimensional coordinate reveals that most variants fall into one of approximately three groups (**Fig. 6A**). Using a density-based hierarchical clustering analysis (HDBSCAN), we unambiguously assigned 182 variants into three distinct clusters based on their expression profiles alone (**Fig. 6A**, Table S4) (5). An analysis of the structural context of these mutants reveals that one of the largest clusters (Cluster 3, n = 76) primarily consists of mutations of soluble loop residues that have little impact on expression and exhibit minimal epistasis (**Fig. 6 A-C**). This analysis also identified a smaller cluster of loop mutations (Cluster 2, n = 30) that moderately decrease expression and exhibit a greater propensity to form epistatic interactions with V276T and/ or W107A relative to the neutral loop mutations in Cluster 3 (**Fig. 6 A-C**). Nevertheless, most of the mutations that exhibit strong positive epistatic interactions with W107A fell into Cluster 1 (n = 76), which primarily consists of mutations within TMDs that significantly decrease expression in all three genetic backgrounds (**Fig. 6 A-C**). A comparison across these clusters suggests positive epistatic interactions with W107A primarily stems from the fact that mutations that compromise expression can only cause a relatively modest reduction in the PME of W107A GnRHR (**Fig. 6 B & C**), which is already predominantly misfolded. In contrast, the moderately disruptive loop mutations in Cluster 2 tend to exhibit more pronounced negative epistatic interactions with V276T than do the highly disruptive TMD mutations in Cluster 1 (**Fig. 6C**). The divergent epistatic interactions that disruptive loop and TMD mutations form with V276T potentially arise from differences in the mechanistic basis for the destabilization caused by these two classes of mutations (see *Discussion*).

Many epistatic interactions arise from the additive energetic effects of mutations on protein folding (26, 40). To determine how these trends relate to the effects of mutations on thermodynamic stability, we utilized Rosetta CM to generate structural models of 243 of the 251 variants that were scored in all three backgrounds and fall within the structured regions of the receptor (35). We then utilized a membrane-optimized Rosetta energy scoring function to approximate the change in the energy of the native structural ensemble of each variant (1). The cluster of loop mutations with lower expression (Cluster 2, avg. $\Delta\Delta G$ = + 2.2 REU) generally contains more destabilizing mutations relative to the cluster of loop mutations with WT-like expression (Cluster 3, avg. $\Delta\Delta G$ = + 0.8 REU), which affirms the general relationship between stability and PME in this system (p = 0.003, **Fig. 6 B & D**). However, the cluster of TMD variants (Cluster 1, avg. $\Delta\Delta G$ = + 3.4 REU) is generally more destabilizing than either of the loop clusters (**Fig. 6D**). A comparison across the clusters suggests that the degree of destabilization generally tracks with the magnitude of positive epistatic interactions that form with W107A, regardless of

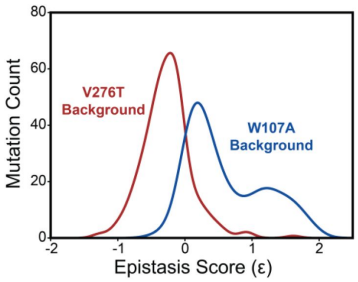


Figure 5.

Epistasis in the mGnRHR double mutant libraries.

A histogram depicts the distribution of epistasis scores (ε) for interactions the subset of 251 high quality secondary mutations form with V276T (red) and W107A (blue).

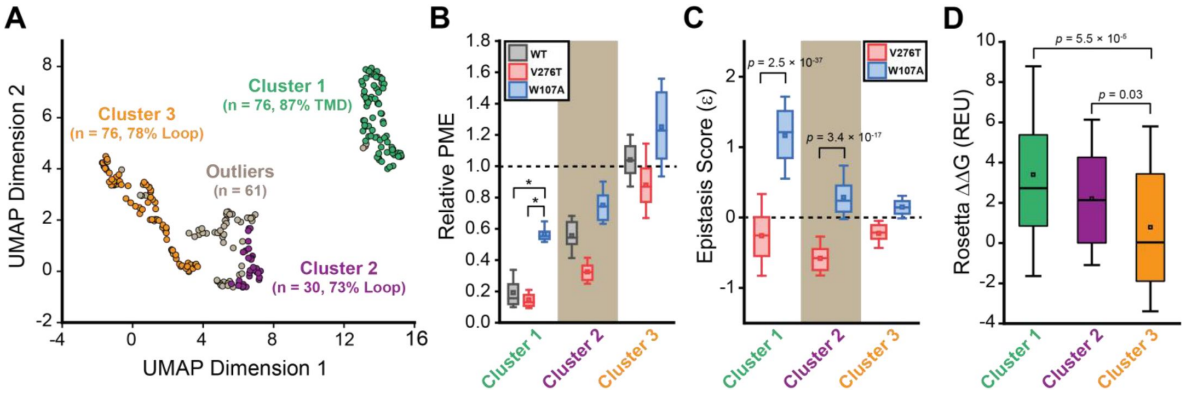


Figure 6.

General trends in mGnRHR epistasis.

Trends associated with the observed pairwise epistasis within mGnRHR are identified using unsupervised learning. A) Uniform manifold approximation projection (UMAP) was used to differentiate variants based on differences in their relative expression in the V276T, W107A, or WT background. Variants are projected onto an arbitrary two-dimensional coordinate based on the results and are colored according to whether they were assigned to cluster 1 (green), cluster 2 (purple), cluster 3 (orange), or were designated as outliers (gray) by HDBSCAN. The percentage of the mutations that fall within TMDs or loops are shown for reference. B) A box and whisker plot depicts the statistical distributions of relative PME values among variants within each cluster in the context of V276T (red), W107A (blue), or WT (Gray) mGnRHR. Select clusters of variants that exhibit statistically different expression profiles according to a Mann-Whitney U-test are indicated (*, $p < 0.001$). A value of 1 corresponds to mutations that have no effect on the PME of mGnRHR in the indicated genetic background. C) A box and whisker plot depicts the statistical distribution of epistasis scores associated with the interactions between the mutations within each cluster and either V276T (red) or W107A (blue). *P*-values for select Mann-Whitney U-tests comparing the interactions of these mutations with V276T and W107A are indicated. A value of 0 indicates that the effects of the two mutations are additive. D) A box and whisker plot depicts the statistical distribution of Rosetta ΔΔG values among mutations within each cluster. *P*-values for select Mann-Whitney U-tests comparing the ΔΔG values across clusters are shown for reference. For panels B-D, the edges of the boxes correspond to the 75th and 25th percentile values while the whiskers reflect the values of the 90th and 10th percentile. The central hash and square within the box represent the average and median values, respectively. These analyses were carried out on a subset of 243 variants with high-quality expression measurements with calculable Rosetta ΔΔG values.

structural context (**Fig. 6 C & D**). This uptick in positive epistasis among mutations that destabilize the native fold and severely reduce PME likely reflects the attenuated effect of destabilizing mutations as the equilibrium fraction of folded protein approaches zero (see *Discussion*, **Fig. 7**). In contrast to W107A, epistatic interactions with V276T are predominantly negative and are most pronounced among moderately destabilizing loop mutations (**Fig. 6 C & D**). The abundance of negative interactions with V276T parallels the observed pairwise epistasis in other folded proteins (26), and likely reflects the higher baseline expression and/ or stability of this variant relative to W107A (**Figs. 2** & **7**). Nevertheless, stability-mediated epistasis cannot account for the attenuation of the epistatic interactions that V276T forms with the highly destabilizing mutations within Cluster 1 relative to the moderately destabilizing mutations within Cluster 2 (**Fig. 6 C & D**). These observations suggest that the magnitude of the epistatic interactions formed by V276T is dependent on not only the change in stability, but also the topological context of the secondary mutation.

Discussion

Many epistatic interactions between mutations within genes encoding soluble proteins arise from their cumulative effects on protein folding energetics (24, 26). Folded proteins can only tolerate a limited number of destabilizing mutations before their cumulative energetic effects cause a cooperative decrease in the fraction of folded protein (**Fig. 7A**). Although this general principle should apply to all proteins, it is unclear how the distinct energetic processes involved in membrane protein biosynthesis and folding may alter such evolutionary couplings. In this study, we utilize DMS to compare the epistatic interactions formed by two mutations that reduce the expression of a misfolding-prone GPCR by selectively destabilizing either its cotranslational membrane integration (V276T) or its native tertiary structure (W107A). We note that the introduction of the V276T mutation into the mouse receptor utilized herein mimics one of the key destabilizing modifications that reduces the expression of hGnRHR relative to mGnRHR (6). Through a focused analysis of 753 total single and double mutants, we find that these mutations give rise to divergent epistatic interactions that modify the PME of mGnRHR (**Fig. 5A**). Mutations that form epistatic interactions with V276T tend to synergistically decrease mGnRHR expression (**Figs. 3 A & B**, **4B**, and **5A**)-a typical epistatic trend that likely arises from the combined destabilizing effects of both mutations. Surprisingly, these interactions are most prominent among moderately destabilizing loop mutations (**Fig. 6C**). In contrast, many of these same secondary mutations instead form positive epistatic interactions with W107A mGnRHR, regardless of their topological context (**Figs. 5A** & **6C**). As this comparison reflects the epistatic interactions formed by an identical set of secondary mutations, we conclude that these context-dependent epistatic interactions arise from differences in the magnitude and/ or the mechanism of the mGnRHR destabilization caused by the V276T and W107A mutations.

In the context of stable proteins, stability-mediated epistasis tends to magnify the effects of deleterious mutations due to the cumulative destabilization that arises from sequential random substitutions (39). This is a fundamental consequence of the cooperative dependence of the fraction of folded protein on the free energy of unfolding. Indeed, most of the random mGnRHR mutations characterized herein are predicted to destabilize the native structure in a manner that coincides with a measurable decrease in the PME in the context of all three genetic backgrounds (**Figs. 4A** & **6D**). Whereas these destabilizing mutations exhibit the expected negative epistatic couplings with V276T, their interactions with W107A skew positive (**Fig. 5A**). This anomalous positive epistasis is potentially a byproduct of the severe destabilizing effects of W107A relative to V276T (**Fig. 2**). While destabilizing secondary mutations can cause significant decreases in the yield of folded WT or V276T mGnRHR, the W107A variant already exhibits marginal surface immunostaining. In the context of this predominantly misfolded, poorly expressed variant, even highly destabilizing secondary mutations can only cause a relatively small decrease in the yield of folded protein. This observation can be more generally related to the

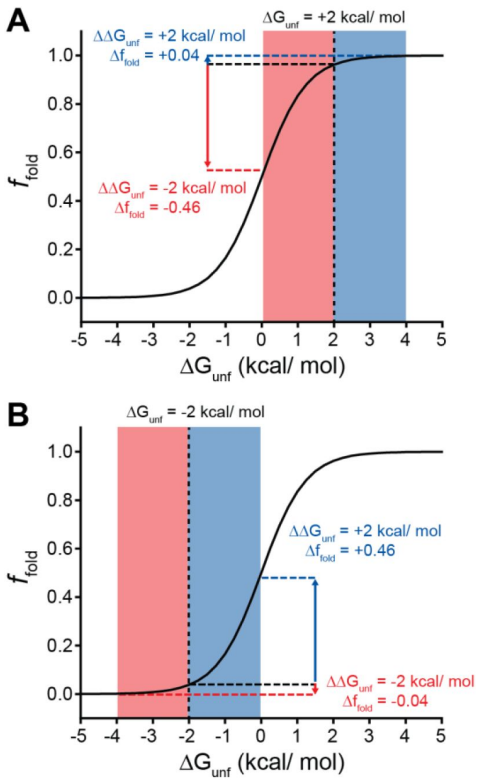


Figure 7.

Context-dependent impacts of stabilizing and destabilizing mutations.

Line plots depict the relationship between the equilibrium fraction of folded protein (f_{fold}) and the free energy of unfolding (ΔG_{unf}). A) A moderately stable protein (ΔG_{unf} = +2 kcal/mol) exhibits a greater change in f_{fold} after acquiring one or more destabilizing mutations (total $\Delta\Delta G_{unf}$ = -2 kcal/mol, red) than after acquiring one or more mutations that stabilize the protein to the same extent (total $\Delta\Delta G_{unf}$ = +2 kcal/mol, blue). B) A moderately unstable protein (ΔG_{unf} = -2 kcal/mol) exhibits a greater change in f_{fold} after acquiring one or more stabilizing mutations (total $\Delta\Delta G_{unf}$ = +2 kcal/mol, blue) than after acquiring one or more mutations that destabilize the protein to the same extent (total $\Delta\Delta G_{unf}$ = -2 kcal/mol, red).

relatively shallow dependence of the fraction of folded protein on the free energy of folding under conditions where folding becomes unfavorable-unfolding transitions become shallow as the fraction of folded protein approaches zero (**Fig. 7B**). Stabilizing mutations can also generate potent positive epistatic couplings in this context given that even small increases in stability can push the system into the steepest portion of the transition zone (**Fig. 7B**). Paradoxically, the asymmetry of the folding curve in this regime suggests destabilizing mutations are more tolerable in the context of an unstable protein-folding efficiency effectively has nowhere to go but upwards. Essentially, the nature of this sigmoidal transition suggests the typical trends associated with stability-mediated epistasis should become inverted as stability decreases. Such considerations could potentially expand the accessible sequence space in metastable proteins, which constitute a significant portion of the proteome (9). Combinations of mutations that would not generally be tolerated in stable proteins may also occur through such mechanism in the context of the numerous misfolded secondary alleles that are carried within the genomes of humans and other diploids with no appreciable fitness cost (e.g. the ΔF508 variant of *CFTR*).

Unlike W107A, the epistatic interactions formed by V276T appear to depend on their topological context-negative epistasis is more prevalent among moderately destabilizing loop variants than highly destabilizing TMD variants (**Fig. 6 C & D**). This observation suggests stability-mediated epistasis between mutations that destabilize the native tertiary structure of integral membrane proteins (Stage II folding) may be distinct from the interactions that arise from the disruption of their translocon-mediated cotranslational membrane integration (Stage I folding). This divergence could potentially reflect the distinction between the energetic principles of protein folding relative to those that guide the insertion of proteins into membranes, which is fundamentally driven by a membrane depth-dependent solvation energetics (13, 23). Counterintuitively, polar mutations that disrupt the membrane integration of individual TMDs also have the potential to form interhelical hydrogen bonds that can stabilize the topological orientation of neighboring TMDs, which can give rise to long-range topological couplings. (11, 22, 25). The propensity of such contacts to form may in some cases depend on the length and hydrophobicity of the loops that connect them (45). Finally, we should note that complex epistatic interactions between residues within TMDs could potentially arise through their propensity to remodel the translocon complex-polar residues within TMDs may affect the recruitment of secondary insertases and/ or intramembrane chaperones such as TRAP, the PAT complex, and the ER membrane protein complex (8, 18, 29, 32, 34, 37). Additional insights into the mechanistic basis for the recruitment of such complexes may therefore be needed to fully rationalize evolutionary couplings between TMDs. We note that, more generally, we suspect that synergistic genetic interactions between mutations could potentially arise from changes in the energetics of many other biosynthetic processes such as cofactor binding, oligomer formation, and chaperone interactions.

Together, our results reveal divergent, mechanism-dependent patterns of pairwise epistasis in the context of a misfolding-prone integral membrane protein. These findings provide novel insights as to how the unique biosynthetic constraints of integral membrane proteins can potentially lead to unique evolutionary patterns. Moreover, our results generally suggest that the deleterious effects of destabilizing mutations are partially attenuated in the context of unstable proteins. Additional investigations are needed to explore how stability-mediated epistasis is modified by other proteostatic constraints, such as the interactions of nascent proteins with cofactors and/ or molecular chaperones.

Materials & Methods

Plasmid Preparation and Characterization of Genetic Libraries

The expression of transiently expressed single mutants was carried out using a previously described pcDNA5 FRT expression vector containing mGnRHR cDNA with an N-terminal hemagglutinin (HA) epitope and bicistronic eGFP (**Fig. 2**) (6). Mutations were introduced by site-directed mutagenesis with PrimeSTAR HS DNA Polymerase (Takara Bio, Shiga, Japan). Biochemical measurements of the membrane integration of TMD6 were carried out using a previously described pGEM vector containing modified leader peptidase (Lep), with TMDs of interest cloned into the H-segment (Fig. S1) (6, 13). Mutations were introduced into the H-segment by site-directed mutagenesis with PrimeSTAR HS DNA Polymerase (Takara Bio, Shiga, Japan). Plasmids were purified using a ZymoPURE Midiprep Kit (Zymo Research, Irvine, CA). DMS experiments were carried out using a previously described pcDNA5 FRT vector containing mGnRHR cDNA bearing an N-terminal influenza hemaglutinin (HA) epitope, followed by an internal ribosome entry site (IRES) and eGFP sequence, which was further modified to be compatible with recombination-based DMS approaches (6). First, an attB recombination site was inserted in place of the CMV promoter by InFusion HD Cloning (Takara Bio, Shiga, Japan). To facilitate the creation of these libraries using nicking mutagenesis, a BbvCI restriction site was introduced by site-directed mutagenesis using PrimeSTAR HS DNA Polymerase (Takara Bio, Shiga, Japan). The V276T and W107A mutations were also introduced by site-directed mutagenesis with PrimeSTAR HS DNA Polymerase (Takara Bio, Shiga, Japan).

To generate mutational libraries, a 10N unique molecular identifier (UMI), or “barcode,” was first inserted into the plasmid backbone using a previously described nicking mutagenesis method (46). Nicking mutagenesis plasmid products were transformed into NEB 10-beta electrocompetent cells (New England Biolabs, Ipswich, MA) and purified using a ZymoPURE Midiprep Kit (Zymo Research, Irvine, CA). A programmed oligo pool (Agilent, Santa Clara, CA) encoding all 19 amino acid substitutions at 85 residues throughout the mGnRHR sequence was then used to generate pools of single and double mutants in the context of WT, V276T, and W107A mGnRHR cDNA using a previously described nicking mutagenesis method (21). The resulting plasmid products were transformed into NEB 10-beta electrocompetent cells (New England Biolabs, Ipswich, MA) and purified using a ZymoPURE Midiprep Kit (Zymo Research, Irvine, CA). To limit the number of mGnRHR variants per UMI, these libraries were then bottlenecked through an additional transformation in NEB Turbo electrocompetent cells (New England Biolabs, Ipswich, MA), and were again purified using a ZymoPURE Midiprep Kit (Zymo Research, Irvine, CA). The resulting plasmid preparations contained detectable levels of a concatemer product, which were removed from the mutational libraries by gel purification using a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Purified products were again transformed into either NEB 10-beta (New England Biolabs, Ipswich, MA) or SIG10 (Sigma-Aldrich, St. Louis, MO) electrocompetent cells, and purified with a ZymoPURE Midiprep Kit (Zymo Research, Irvine, CA).

To associate each UMI with its corresponding GnRHR variant, the final plasmid libraries were also sequenced using PacBio SMRT sequencing. Briefly, plasmid libraries were first double-digested with PmlI and MfeI-HF restriction enzymes (New England Biolabs, Ipswich, MA), then purified using a Zymo DNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA). Complete digestion was verified by analyzing the products on an agarose gel, purity was assessed with a Synergy Neo2 microplate reader (BioTek, Winooski, VT), and the final yield was quantified using a Qubit4 fluorometer (ThermoFisher, Waltham, MA). The 1537 bp fragment for each library, containing the 10N UMI and GnRHR open reading frame, was isolated on a BluePippin instrument (Sage Science, Beverly, MA). These fragments were then assembled into PacBio libraries and sequenced on the Sequel II system with a 30-hour runtime.

The sequence of each plasmid-derived fragment containing mGnRHR variants and their corresponding UMIs was determined by reconstructing the circular consensus sequences (CCS) for each independent well in the PacBio readout. mGnRHR libraries were sequenced to a depth of 1,825,010 wells for WT, 1,545,151 wells for V276T, and 1,497,359 wells for W107A. A minimum coverage of five passes around each circular fragment (subread) was required for incorporation of sequencing data from an individual well into the analysis. CCS were mapped to the reference open reading frame using minimap2. CCS that did not contain a complete 10-base UMI sequence or fully cover the target ORF were excluded from the analysis. Variations in the mGnRHR sequence of individual fragments were called using SAMtools. Reads were then grouped according to their UMI sequences. Given that insertion and deletion (indel) artifacts are prevalent within PacBio sequencing data, we only removed UMIs from the analysis if indels were detected at a specific position within the UMI in at least 10% of the subreads. Indels observed within the open reading frame were only called if their occurrence was judged to be statistically significant relative to the background indel rate for the reads according to a binomial test (p <= .05). Codon substitutions encoded in our primer library (Table S3) were called within the mGnRHR open reading frame only if they were detected in at least 75% of the individual subreads. Only UMIs that were found to be associated with a single mGnRHR variant were included in the final analysis. Our final "dictionaries" for the WT, V276T, and W107A libraries contained 3,261, 531, and 4,962 unique UMIs, respectively, that could be indexed to a corresponding mGnRHR variant.

Preparation of Cellular Libraries and Cell Sorting

A previously described HEK293T cell line with a genomic Tet-Bxb1-BFP landing pad was obtained from D. Fowler (18). Cells were grown at 37 oC and 5.0% CO_2 in complete media, made of Dulbecco's Modified Eagle's medium (DMEM, Gibco, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS, Gibco, Grand Island, NY), 0.5% penicillin (Gibco, Grand Island, NY), and 0.5% streptomycin (Gibco, Grand Island, NY). 2 million cells were plated in 10-cm tissue culture plates and co-transfected the next day with 475 ng of plasmid encoding bxb1 recombinase and 7125 ng of the plasmid library using Lipofectamine 3000 (Invitrogen, Carlsbad, CA). Cells were grown at 33 °C for 4 days after transfection. 2 days after transfection, the cells were induced with 2 µg/mL doxycycline. All cells were expanded into 15-cm tissue culture plates 6 days after transfection, and 10 million cells were re-plated into new 15-cm tissue culture plates 3 days later.

Twelve days after transfection, the cells were washed with 1X phosphate-buffered saline (PBS, Gibco, Grand Island, NY), and harvested with TrypLE Express protease (Gibco, Grand Island, NY). Cells were washed twice with 2% fetal bovine serum in PBS (wash buffer), resuspended in PBS containing 2% FBS to 10M cells/ mL, and then sorted on a BD FACS Aria II (BD Biosciences, Franklin Lakes, NJ). Recombined cells were isolated based on their characteristic gain of bicistronic eGFP and loss of BFP expression (18). Forward and side scatter profiles were first used to gate for intact live cells, and cells with eGFP-positive (488 nm laser, 530/30 nm emission filter) and BFP-negative (405 nm laser, 450/50 nm emission filter) fluorescence profiles were collected in complete media supplemented with 10% FBS (Gibco, Grand Island, NY) and then plated in 10-cm tissue culture plates. Cells were induced with 2 µg/mL doxycycline 24 hours after sorting. 4 days after sorting, 10 million cells were passaged into 15-cm tissue culture plates.

7 days after sorting for eGFP-positive cells, cells were sorted again according to the surface expression of GnRHR variants. Briefly, cells were first washed with 1X phosphate-buffered saline (PBS, Gibco, Grand Island, NY), and harvested with TrypLE Express protease (Gibco, Grand Island, NY). GnRHRs expressed at the cell surface were immunostained for 30 minutes in the dark with DyLight 550-conjugated anti-HA antibody (Invitrogen, Carlsbad, CA). Cells were then washed twice with PBS containing 2% FBS, and sorted on a BD FACS Aria II (BD Biosciences, Franklin Lakes, NJ). Forward and side scatter profiles were used to gate for intact live cells, and eGFP fluorescence was used to gate for recombined cells. Cells were then sorted into even quartiles based on the amount of DyLight 550 fluorescence (561 nm laser, 585/15 nm emission filter). Sorted cells were collected in complete media supplemented with 10% FBS (Gibco, Grand Island, NY) and plated in 10-cm

tissue culture plates. Cells were allowed to grow to confluency and then were washed with 1X phosphate-buffered saline (PBS, Gibco, Grand Island, NY) and harvested with TrypLE Express protease (Gibco, Grand Island, NY). Cell pellets were frozen for subsequent analysis.

Extraction of Genomic DNA and Next Generation Sequencing

Genomic DNA (gDNA) was extracted from cell pellets using the Sigma GenElute Mammalian gDNA Miniprep Kit (Sigma-Aldrich, St. Louis, MO). DNA amplicons for Illumina sequencing were produced from the gDNA using a previously described semi-nested polymerase chain reaction (PCR) technique (18). In order to selectively amplify recombined DNA, the first PCR utilized a primer which anneals upstream of the 10N UMI and a primer which anneals in the BFP-encoding region of the landing pad. Eight replicate PCRs were carried out with 2.5 µg of gDNA and HiFi HotStart Ready Mix (Kapa Biosystems, Wilmington, MA). These PCRs were limited to 7 cycles to avoid PCR bias. PCR products were purified using a ZR-96 DNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA), and replicate products were then pooled. 10 µL of this first PCR product was then used as the template for a second PCR, which utilized primers that incorporate Illumina adapter sequences to both ends of the 10N UMI. To avoid PCR bias, these reactions were monitored by real-time PCR on a StepOne RT-PCR instrument (Applied Biosystems, Waltham, MA) and terminated during mid-log amplification (17 or 18 cycles). Four replicate PCRs were carried out with HiFi HotStart Ready Mix (Kapa Biosystems, Wilmington, MA) and SYBR Green fluorescent dye (Applied Biosystems, Waltham, MA). Replicate PCR products were then combined and purified with a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). To further improve the yield and quality of the amplicons, an additional 6-cycle PCR was carried out with the HiFi HotStart Ready Mix (Kapa Biosystems, Wilmington, MA) and primers that preserve the DNA sequence. These PCR products were then purified using a Zymo DNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA). Finally, amplicon yield and quality were assessed using an Agilent 2200 TapeStation with D1000 tape (Agilent Technologies, Santa Clara, CA) and a Synergy Neo2 microplate reader (BioTek, Winooski, VT). Amplicons were sequenced using an Illumina NextSeq150 flow cell, paired end, with 50% PhiX content.

Plasma Membrane Expression and Epistasis Measurements

Illumina sequencing of the recombined UMIs within each sorted cell fraction were matched to the corresponding GnRHR mutants using the PacBio sequencing data and subjected to previously described quality filtering (27). Estimations of PME for each mutant were then calculated as previously described (27). PME measurements were averaged across two biological replicates. Within each library, the average fluorescence intensity values of the mutants were ranked and converted into a percentile. Mutants which differed in their percentile values by more than 25% across the two replicates were excluded from analysis. The use of percentile-based filtering overcomes intrinsic limitations associated with the distinct dynamic range constraints that occur because of the distinct surface immunostaining profiles of these three cellular libraries (**Fig. 3**). Intensity values that meet these criteria are highly correlated across the two biological replicates (average Spearman's ρ=0.936 across the three libraries, Supplementary Figs. S3 & S4). To facilitate the comparison of epistatic interactions within a common set of secondary mutations, we excluded mutants that did not pass quality filtering in all three libraries from our analysis. After data filtering, we obtained PME measurements for 251 mutants in all three libraries (Table S 4-6). These PME measurements were then utilized to calculate epistasis scores for the double mutants in the V276T and W107A libraries using the following equation:

$$\varepsilon = \ln(\omega_{ab}) - \ln(\omega_a) - \ln(\omega_b)$$

where ε is the epistasis score, ω_{ab} is the fluorescence intensity of the double mutant relative to WT mGnRHR, ω_a is the fluorescence intensity of the single mutant in the WT library relative to WT mGnRHR, and ω_b is the fluorescence intensity of the background mutation (either V276T or W107A) relative to WT mGnRHR, as previously described (26).

Expression Measurements of Individual GnRHR Variants

HEK283T cells were obtained from ATCC and grown under the same conditions described above. GnRHR variants were transiently expressed in these cells, and the plasma membrane and intracellular expression of these variants was measured by flow cytometry, as previously described (6). The reported expression levels represent measurements from three biological replicates.

In vitro Translation of Chimeric Lep Proteins

Messenger RNA (mRNA) of chimeric Lep proteins was generated and used for *in vitro* translation reactions as previously described (6). Translation products were then analyzed by SDS-PAGE as previously described (6). The intensities of singly (*G1*) and doubly (*G2*) glycosylated protein were quantified by densitometry in ImageJ software. Apparent transfer free energy values representing transfer of the TMDs of interest from the translocon to the membrane were then calculated using the following equation:

$$\Delta G_{app} = -RTln(K_{app}) = -RTln(\frac{G1}{G2})$$

where ΔG_{app} is the free energy for transfer of the TMD into the membrane, R is the universal gas constant, T is the temperature, K_{app} is the equilibrium constant for transfer of the TMD into the membrane, *G1* is the intensity of the singly glycosylated band, and *G2* is the intensity of the doubly glycosylated band, as previously described (12). The reported transfer free energy values represent the average of three experimental replicates.

Structural Modeling of M. musculus GnRHR

The sequence of mGnRHR was obtained from Uniprot (accession number Q01776) and a homology model in the inactive state was generated as previously described (6). Images were generated by structurally aligning this mGnRHR model to an experimentally determined inactive-state human GnRHR crystal structure (GNRHR, PDB 7BR3, 2.79 Å) using the Super command in PyMol (Schrödinger, Inc., New York, NY) (47).

Computational Estimates for the Stability of mGnRHR Variants

Structural estimates for the effects of mutations on the stability of mGnRHR were determined using a previously described Rosetta based protocol featuring a membrane protein optimized energy function (1, 2). Briefly, the homology model of mGnRHR described above was used as the starting structure for computational stability estimates. A spanfile describing the transmembrane regions was created for mGnRHR using OCTOPUS predictions from Topcons (3, 42). The homology model was transformed into membrane coordinates using the mp_transform application. An ensemble of structures of each variant and wildtype (50 iterations) was generated from the homology model. For variants containing two mutations, the mutations were both introduced at the same time. The ΔΔG for each variant was calculated by subtracting the average score of the three lowest scoring variant models from the average score of the three lowest scoring WT models ($\Delta\Delta G = \Delta G_{mut} - \Delta G_{wt}$).

Unsupervised Learning

To classify mGnRHR variants based on their expression profiles, we first standardized their scores using the Standard Scaler tool from Python's scikit-learn library in order to ensure that the clustering algorithms are not affected by the scale of the data. Next, we then reduced the dimensionality of the standardized data using UMAP (20). We then used a density-based clustering algorithm (HDBSCAN) to cluster variants based on their expression profiles (19). Unsupervised learning analyses were carried out based on deep mutational scanning measurements alone, and did not incorporate any structural and/ or energetic parameters.

Quantification and Analysis

Flow cytometry data were analyzed in FlowJo software (Treestar, Ashland, OR), and *in vitro* translation data were analyzed in ImageJ software. PME measurements were analyzed in Origin software (OriginLab, Northampton, MA) and mapped onto the mGnRHR homology model in PyMol (Schrödinger, Inc., New York, NY).

Data and Resource Availability

Code for the analysis of sequencing data for deep mutational scanning experiments can be found on the Schlebach Lab GitHub page (*https://github.com/schebachlab*). Fluorescence intensity values for mutants in all three mutational libraries are included as Excel files in the Supplementary Materials. Additional data is available upon request.

Acknowledgements

We thank Christiane Hassel and the Indiana University Bloomington Flow Cytometry Core Facility, as well as Luke Tallon, Xuechu Zhao, Lisa Sadzewicz, and the University of Maryland Institute for Genome Sciences for their experimental support. This work was supported by a National Science Foundation (NSF) Graduate Research Fellowship (1342962 to L. M. C.) and by the National Institutes of Health (R01GM129261 to J.P.S). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

References

1. Alford RF, Koehler Leman J, Weitzner BD, Duran AM, Tilley DC, et al. (2015) **An Integrated Framework Advancing Membrane Protein Modeling and Design** *PLoS Comput Biol* **11**

2. Barlow KA, Ó Conchúir S, Thompson S, Suresh P, Lucas JE, et al. (2018) **Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein–Protein Binding Affinity upon Mutation** *J Phys Chem B* **122**:5389–99

3. Bernsel A, Viklund H, Hennerdal A, Elofsson A. (2009) **TOPCONS: consensus prediction of membrane protein topology** *Nucleic Acids Res* **37**:W465–8

4. Bloom JD, Raval A, Wilke CO (2007) **Thermodynamics of neutral protein evolution** *Genetics* **175**:255–66

5. Campello RJGB, Moulavi D, Sander J., Pei J, Tseng VS, Cao L, Xu G, Moroda H (2013) **Density-Based Clustering Based on Hierarchical Density Estimates** *In Advances in Knowledge Discovery and Data Mining* :160–72

6. Chamness LM, Zelt NB, Harrington HR, Kuntz CP, Bender BJ, et al. (2021) **Molecular basis for the evolved instability of a human G-protein coupled receptor** *Cell Rep* **37**

7. Faber MS, Wrenbeck EE, Azouz LR, Steiner PJ, Whitehead TA (2019) **Impact of in Vivo Protein Folding Probability on Local Fitness Landscapes** *Mol Biol Evol* **36**:2764–77

8. Gemmer M, Chaillet ML, van Loenhout J, Cuevas Arenas R, Vismpas D, et al. (2023) **Visualization of translation and protein biogenesis at the ER membrane** *Nature* **614**:160–67

9. Ghosh K, Dill K. (2010) **Cellular proteomes have broad distributions of protein stability** *Biophys J* **99**:3996–4002

10. Gong LI, Suchard MA, Bloom JD (2013) **Stability-mediated epistasis constrains the evolution of an influenza protein** *Elife* **2013**:1–19

11. Hermansson M, von Heijne G. (2003) **Inter-helical hydrogen bond formation during membrane protein integration into the ER membrane** *J Mol Biol* **334**:803–9

12. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, et al. (2005) **Recognition of transmembrane helices by the endoplasmic reticulum translocon** *Nature* **433**

13. Hessa T, Meindl-Beinker NM, Bernsel A, Kim H, Sato Y, et al. (2007) **Molecular code for transmembrane-helix recognition by the Sec61 translocon** *Nature* **450**

14. Janovick JA, David Stewart M, Jacob D, Martin LD, Deng JM, et al. (2013) **Restoration of testis function in hypogonadotropic hypogonadal mice harboring a misfolded GnRHR mutant by pharmacoperone drug therapy** *PNAS* **110**:21030–35

15. Janovick JA, Knollman PE, Brothers SP, Ayala-Yáñ Ez R, Aziz AS, Conn PM (2006) **Regulation of G Protein-coupled Receptor Trafficking by Inefficient Plasma Membrane Expression MOLECULAR BASIS OF AN EVOLVED STRATEGY** *J Biol Chem* **281**:8417–25

16. Jones EM, Lubock NB, Venkatakrishnan AJ, Wang J, Tseng AM, et al. (2020) **Structural and functional characterization of G protein-coupled receptors with deep mutational scanning** *Elife* **9**:1–52

17. Marinko JT, Huang H, Penn WD, Capra JA, Schlebach JP, Sanders CR (2019) **Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis** *Chem Rev* **119**:5537–5606

18. Matreyek KA, Stephany JJ, Fowler DM (2017) **A platform for functional assessment of large variant libraries in mammalian cells** *Nucleic Acids Res* **45**:1–12

19. McInnes L, Healy J, Astels S. (2017) **hdbscan: Hierarchical density based clustering** *The Journal of Open Source Software* **2**

20. McInnes L, Healy J, Saul N, Großberger L. (2018) **UMAP: Uniform Manifold Approximation and Projection** *J Open Source Softw* **3**

21. Medina-Cucurella A v., Steiner PJ, Faber MS, Beltrán J, Borelli AN, et al. (2019) **User-defined single pot mutagenesis using unamplified oligo pools** *Protein Engineering, Design and Selection* **32**:41–45

22. Meindl-Beinker NM, Lundin C, Nilsson I, White SH, von Heijne G. (2006) **Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly** *EMBO Rep* **7**:1111–16

23. Moon CP, Fleming KG (2011) **Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers** *PNAS* **108**:10174–77

24. Nedrud D, Coyote-Maestas W, Schmidt D. (2021) **A large-scale survey of pairwise epistasis reveals a mechanism for evolutionary expansion and specialization of PDZ domains** *Proteins* **89**:899–914

25. Öjemalm K, Halling KK, Nilsson IM, von Heijne G. (2012) **Orientational Preferences of Neighboring Helices Can Drive ER Insertion of a Marginally Hydrophobic Transmembrane Helix** *Mol Cell* **45**:529–40

26. Olson CA, Wu NC, Sun R. (2014) **A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain** *Current Biology* **24**:2643–51

27. Penn WD, McKee AG, Kuntz CP, Woods H, Nash V, et al. (2020) **Probing Biophysical Sequence Constraints within the Transmembrane Domains of Rhodopsin by Deep Mutational Scanning** *Sci Adv* **6**

28. Pfeffer S, Dudek J, Zimmermann R, Förster F. (2016) **Organization of the native ribosome–translocon complex at the mammalian endoplasmic reticulum membrane** *Biochim Biophys Acta Gen Subj* **1860**:2122–29

29. Pleiner T, Tomaleri GP, Januszyk K, Inglis AJ, Hazu M, Voorhees RM (2020) **Structural basis for membrane insertion by the human ER membrane protein complex** *Science* **369**:433–36

30. Popot JL, Engelman DM (1990) **Membrane Protein Folding and Oligomerization** *Biochemistry* **29**:4031–37

31. Schlebach JP, Sanders CR (2016) **The Safety Dance: Biophysics of Membrane Protein Folding and Misfolding in a Cellular Context** *Q Rev Biophys* **48**:1–34

32. Shurtleff MJ, Itzhak DN, Hussmann JA, Schirle Oakdale NT, Costa EA, et al. (2018) **The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins** *Elife* **7**

33. Smalinskaite L, Hegde RS (2022) **The Biogenesis of Multipass Membrane Proteins** *Cold Springs Harbor Perspectives in Biology*

34. Smalinskaitė L, Kim MK, Lewis AJO, Keenan RJ, Hegde RS (2022) **Mechanism of an intramembrane chaperone for multipass membrane proteins** *Nature* **611**

35. Song Y, DiMaio F, Yu-Ruei Wang R, Kim D, Miles C, et al. (2013) **High resolution comparative modeling with RosettaCM** *Structure* **21**

36. Starr TN, Thornton JW (2016) **Epistasis in protein evolution** *Protein Science* **25**:1204–18

37. Sundaram A, Yamsek M, Zhong F, Hooda Y, Hegde RS, Keenan RJ (2022) **Substrate-driven assembly of a translocon for multipass membrane proteins** *Nature* **611**:167–72

38. Telenti A, Pierce LCT, Biggs WH, di Iulio J, Wong EHM, et al. (2016) **Deep sequencing of 10,000 human genomes** *PNAS* **113**:11901–6

39. Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS (2007) **The Stability Effects of Protein Mutations Appear to be Universally Distributed** *J Mol Biol* **369**:1318–32

40. Tokuriki N, Tawfik DS (2009) **Stability effects of mutations and protein evolvability** *Curr Opin Struct Biol* **19**:596–604

41. van den Berg B, Clemons WM, Collinson I, Modis Y, Hartmann E, et al. (2003) **X-ray structure of a protein-conducting channel** *Nature* **427**:36–44

42. Viklund H, Elofsson A. (2008) **OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar** *Bioinformatics* **24**:1662–68

43. White SH, Von Heijne G. (2008) **How Translocons Select Transmembrane Helices**

44. Wiseman RL, Koulov A, Powers E, Kelly JW, Balch WE (2007) **Protein energetics in maturation of the early secretory pathway** *Curr Opin Cell Biol* **19**

45. Woodall NB, Hadley S, Yin Y, Bowie JU (2017) **Complete topology inversion can be part of normal membrane protein biogenesis** *Protein Science* **26**:824–33

46. Wrenbeck EE, Klesmith JR, Stapleton JA, Adeniran A, Tyo KEJ, Whitehead TA (2016) **Plasmid-based one-pot saturation mutagenesis** *Nat Methods* **13**:928–30

47. Yan W, Cheng L, Wang W, Wu C, Yang X, et al. (2020) **Structure of the human gonadotropin-releasing hormone receptor GnRH1R reveals an unusual ligand binding mode** *Nat Commun* **11**

Article and author information

Laura M. Chamness
Department of Chemistry, Indiana University, Bloomington, Indiana, USA
ORCID iD: 0000-0002-3971-2799

Charles P. Kuntz
Department of Chemistry, Purdue University, West Lafayette, Indiana, USA

Andrew G. McKee
Department of Chemistry, Indiana University, Bloomington, Indiana, USA

Wesley D. Penn
Department of Chemistry, Indiana University, Bloomington, Indiana, USA

Christopher M. Hemmerich
Center for Genomics and Bioinformatics, Indiana University, Bloomington, Indiana, USA

Douglas B. Rusch
Center for Genomics and Bioinformatics, Indiana University, Bloomington, Indiana, USA

Hope Woods
Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA, Chemical and Physical Biology Program, Vanderbilt University, Nashville, Tennessee, USA

Dyotima
Department of Chemistry, Indiana University, Bloomington, Indiana, USA

Jens Meiler
Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA, Institute for Drug Development, Leipzig University, Leipzig, SAC, Germany
ORCID iD: 0000-0001-8945-193X

Jonathan P. Schlebach
Department of Chemistry, Indiana University, Bloomington, Indiana, USA, Department of Chemistry, Purdue University, West Lafayette, Indiana, USA
For correspondence: jschleba@purdue.edu
ORCID iD: 0000-0003-0955-7633

Copyright

Editors

Reviewing Editor
David Drew
Stockholm University, Stockholm, Sweden

Senior Editor
Kenton Swartz
National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, United States of America

Reviewer #1 (Public Review):

Summary:
The paper carries out an impressive and exhaustive non-sense mutagenesis using deep mutational scanning (DMS) of the gonadotropin-releasing hormone receptor for the WT protein and two single point mutations that I) influence TM insertion (V267T) and ii) influence protein stability (W107A), and then measures the effect of these mutants on correct plasma membrane expression (PME).

Overall, most mutations decreased mGnRHR PME levels in all three backgrounds, indicating poor mutational tolerance under these conditions. The W107A variant wasn't really recoverable with low levels of plasma membrane localisation. For the V267T variant, most additional mutations were more deleterious than WT based on correct trafficking, indicating a synergistic effect. As one might expect, there was a higher degree of positive correlation between V267T/W107A mutants and other mutants located in TM regions, confirming that improper trafficking was a likely consequence of membrane protein co-translational folding. Nevertheless, context is important, as positive synergistic mutants in the V27T could be negative in the W107A background and vice versa. Taken together, this important study highlights the complexity of membrane protein folding in dissecting the mechanism-dependent impact of disease-causing mutations related to improper trafficking.

Strengths:
This is a novel and exhaustive approach to dissecting how receptor mutations under different mutational backgrounds related to co-translational folding, could influence membrane protein trafficking.

Weaknesses:
The premise for the study requires an in-depth understanding of how the single-point mutations analysed affect membrane protein folding, but the single-point mutants used seem to lack proper validation. Furthermore, plasma membrane expression has been used as a proxy for incorrect membrane protein folding, but this not necessarily be the case, as even correctly folded membrane proteins may not be trafficked correctly, at least, under heterologous expression conditions. In addition, mutations can affect trafficking and potential post-translational modifications, like glycosylation.

https://doi.org/10.7554/eLife.92406.1.sa1

Reviewer #2 (Public Review):

Summary:
In this paper, Chamness and colleagues make a pioneering effort to map epistatic interactions among mutations in a membrane protein. They introduce thousands of mutations to the mouse GnRH Receptor (GnRHR), either under wild-type background or two mutant backgrounds, representing mutations that destabilize GnRHR by distinct mechanisms. The first mutant background is W107A, destabilizing the tertiary fold, and the second, V276T, perturbing the efficiency of cotranslational insertion of TM6 to the membrane, which is essential for proper folding. They then measure the surface expression of these three mutant libraries, using it as a proxy for protein stability, since misfolded proteins do not typically make it to the plasma membrane. The resulting dataset is then used to shed light on how diverse mutations interact epistatically with the two genetic background mutations. Their

main conclusion is that epistatic interactions vary depending on the degree of destabilization and the mechanism through which they perturb the protein. The mutation V276T forms primarily negative (aggravating) epistatic interactions with many mutations, as is common to destabilizing mutations in soluble proteins. Surprisingly, W107A forms many positive (alleviating) epistatic interactions with other mutations. They further show that the locations of secondary mutations correlate with the types of epistatic interactions they form with the above two mutants.

Strengths:
Such a high throughput study for epistasis in membrane proteins is pioneering, and the results are indeed illuminating. Examples of interesting findings are that: (1) No single mutation can dramatically rescue the destabilization introduced by W107A. (2) Epistasis with a secondary mutation is strongly influenced by the degree of destabilization introduced by the primary mutation. (3) Misfolding caused by mis-insertion tends to be aggravated by further mutations. The discussion of how protein folding energetics affects epistasis (Fig. 7) makes a lot of sense and lays out an interesting biophysical framework for the findings.

Weaknesses:
The major weakness comes from the potential limitations in the measurements of surface expression of severely misfolded mutants. This point is discussed quite fairly in the paper, in statements like "the W107A variant already exhibits marginal surface immunostaining" and many others. It seems that only about 5% of the W107A makes it to the plasma membrane compared to wild-type (Figures 2 and 3). This might be a low starting point from which to accurately measure the effects of secondary mutations.

Still, the authors claim that measurements of W107A double mutants "still contain cellular subpopulations with surface immunostaining intensities that are well above or below that of the W107A single mutant, which suggests that this fluorescence signal is sensitive enough to detect subtle differences in the PME of these variants". I was not entirely convinced that this was true. Firstly, I think it would be important to test how much noise these measurements have and how much surface immunostaining the W107A mutant displays above the background of cells that do not express the protein at all. But more importantly, it is not clear if under this regimen surface expression still reports on stability/protein fitness. It is unknown if the W107A retains any function or folding at all. For example, it is possible that the low amount of surface protein represents misfolded receptors that escaped the ER quality control. The differential clustering of epistatic mutations (Fig. 6) provides some interesting insights as to the rules that dictate epistasis, but these too are dominated by the magnitude of destabilization caused by one of the mutations. In this case, the secondary mutations that had the most interesting epistasis were exceedingly destabilizing. With this in mind, it is hard to interpret the results that emerge regarding the epistatic interactions of W107A. Furthermore, the most significant positive epistasis is observed when W107A is combined with additional mutations that almost completely abolish surface expression. It is likely that either mutation destabilizes the protein beyond repair. Therefore, what we can learn from the fact that such mutations have positive epistasis is not clear to me. Based on this, I am not sure that another mutation that disrupts the tertiary folding more mildly would not yield different results.

With that said, I believe that the results regarding the epistasis of V276T with other mutations are strong and very interesting on their own.

Additionally, the study draws general conclusions from the characterization of only two mutations, W107A and V276T. At this point, it is hard to know if other mutations that perturb insertion or tertiary folding would behave similarly. This should be emphasized in the text.

Some statistical aspects of the study could be improved:

1. It would be nice to see the level of reproducibility of the biological replicates in a plot, such as scatter or similar, with correlation values that give a sense of the noise level of the measurements. This should be done before filtering out the inconsistent data.

2. The statements "Variants bearing mutations within the C- terminal region (ICL3-TMD6-ECL3-TMD7) fare consistently worse in the V276T background relative to WT (Fig. 4 B & E)." and "In contrast, mutations that are 210 better tolerated in the context of W107A mGnRHR are located 211 throughout the structure but are particularly abundant among residues 212 in the middle of the primary structure that form TMD4, ICL2, and ECL2 213 (Fig. 4 C & F)." are both hard to judge. Inspecting Figures 4B and C does not immediately show these trends, and importantly, a solid statistical test is missing here. In Figures 4E and F the locations of the different loops and TMs are not indicated on the structure, making these statements hard to judge.

3. The following statement lacks a statistical test: "Notably, these 98 variants are enriched with TMD variants (65% TMD) relative to the overall set of 251 variants (45% TMD)." Is this enrichment significant? Further in the same paragraph, the claim that "In contrast to the sparse epistasis that is generally observed between mutations within soluble proteins, these findings suggest a relatively large proportion of random mutations form epistatic interactions in the context of unstable mGnRHR variants". Needs to be backed by relevant data and statistics, or at least a reference.

https://doi.org/10.7554/eLife.92406.1.sa0

Author Response

We are grateful for these balanced, nuanced evaluations of our work concerning the observed epistatic trends and our interpretations of their mechanistic origins. Overall, we think the reviewers have done an excellent job at recognizing the novel aspects of our findings while also discussing the caveats associated with our interpretations of the biophysical effects of these mutations. We believe it is important to consider both of these aspects of our work in order to appreciate these advances and what sorts of pertinent questions remain.

Notably, both reviewers suggest that a lack of experimental approaches to compare the conformational properties of GnRHR variants weakens our claims. We would first humbly suggest that this constitutes a more general caveat that applies to nearly all investigations of the cellular misfolding of α-helical membrane proteins. Whether or not any current in vitro folding measurements report on conformational transitions that are relevant to cellular protein misfolding reactions remains an active area of debate (discussed further below). Nevertheless, while we concede that our structural and/ or computational evaluations of various mutagenic effects remain speculative, prevailing knowledge on the mechanisms of membrane protein folding suggest our mutations of interest (V276T and W107A) are highly unlikely to promote misfolding in precisely the same way. Thus, regardless of whether or not we were able experimentally compare the relevant folding energetics of GnRHR variants, we are confident that the distinct epistatic interactions formed by these mutations reflect variations in the misfolding mechanism and that they are distinct from the interactions that are observed in the context of stable proteins. In the following, we provide detailed considerations concerning these caveats in relation to the reviewers' specific comments.

Response to the Reviews

We are grateful for these balanced, nuanced evaluations of our work concerning the observed epistatic trends and our interpretations of their mechanistic origins. Overall, we think the reviewers have done an excellent job at recognizing the novel aspects of our findings while also discussing the caveats associated with our interpretations of the biophysical effects of these mutations. We believe it is important to consider both of these aspects of our work in order to appreciate these advances and what sorts of pertinent questions remain.

Notably, both reviewers suggest that a lack of experimental approaches to compare the conformational properties of GnRHR variants weakens our claims. We would first humbly suggest that this constitutes a more general caveat that applies to nearly all investigations of the cellular misfolding of α-helical membrane proteins. Whether or not any current in vitro folding measurements report on conformational transitions that are relevant to cellular protein misfolding reactions remains an active area of debate (discussed further below). Nevertheless, while we concede that our structural and/ or computational evaluations of various mutagenic effects remain speculative, prevailing knowledge on the mechanisms of membrane protein folding suggest our mutations of interest (V276T and W107A) are highly unlikely to promote misfolding in precisely the same way. Thus, regardless of whether or not we were able experimentally compare the relevant folding energetics of GnRHR variants, we are confident that the distinct epistatic interactions formed by these mutations reflect variations in the misfolding mechanism and that they are distinct from the interactions that are observed in the context of stable proteins. In the following, we provide detailed considerations concerning these caveats in relation to the reviewers' specific comments.

Reviewer #1 (Public Review):

The paper carries out an impressive and exhaustive non-sense mutagenesis using deep mutational scanning (DMS) of the gonadotropin-releasing hormone receptor for the WT protein and two single point mutations that I) influence TM insertion (V267T) and ii) influence protein stability (W107A), and then measures the effect of these mutants on correct plasma membrane expression (PME).

Overall, most mutations decreased mGnRHR PME levels in all three backgrounds, indicating poor mutational tolerance under these conditions. The W107A variant wasn't really recoverable with low levels of plasma membrane localisation. For the V267T variant, most additional mutations were more deleterious than WT based on correct trafficking, indicating a synergistic effect. As one might expect, there was a higher degree of positive correlation between V267T/W107A mutants and other mutants located in TM regions, confirming that improper trafficking was a likely consequence of membrane protein co-translational folding. Nevertheless, context is important, as positive synergistic mutants in the V27T could be negative in the W107A background and vice versa. Taken together, this important study highlights the complexity of membrane protein folding in dissecting the mechanism-dependent impact of disease-causing mutations related to improper trafficking.

Strengths

This is a novel and exhaustive approach to dissecting how receptor mutations under different mutational backgrounds related to co-translational folding, could influence membrane protein trafficking.

Weaknesses

> *The premise for the study requires an in-depth understanding of how the single-point mutations analysed affect membrane protein folding, but the single-point mutants used seem to lack proper validation.*

Given our limited understanding of the structural properties of misfolded membrane proteins, it is unclear whether the relevant conformational effects of these mutations can be unambiguously validated using current biochemical and/ or biophysical folding assays. X-ray crystallography, cryo-EM, and NMR spectroscopy measurements have demonstrated that many purified GPCRs retain native-like structural ensembles within certain detergent micelles, bicelles, and/ or nanodiscs. However, helical membrane protein folding measurements typically require titration with denaturing detergents to promote the formation of a denatured state ensemble (DSE), which will invariably retain considerable secondary structure. Given that the solvation provided by mixed micelles is clearly distinct from that of native membranes, it remains unclear whether these DSEs represent a reasonable proxy for the misfolded conformations recognized by cellular quality control (QC, see https://doi.org/10.1021/acs.chemrev.8b00532). Thus, the use and interpretation of these systems for such purposes remains contentious in the membrane protein folding community. In addition to this theoretical issue, we are unaware of any instances in which GPCRs have been found to undergo reversible denaturation in vitro- a practical requirement for equilibrium folding measurements (https://doi.org/10.1146/annurev-biophys-051013-022926). We note that, while the resistance of GPCRs to aggregation, proteolysis, and/ or mechanical unfolding have also been probed in micelles, it is again unclear whether the associated thermal, kinetic, and/ or mechanical stability should necessarily correspond to their resistance to cotranslational and/ or posttranslational misfolding. Thus, even if we had attempted to validate the computational folding predictions employed herein, we suspect that any resulting correlations with cellular expression may have justifiably been viewed by many as circumstantial. Simply put, we know very little about the non-native conformations are generally involved in the cellular misfolding of α-helical membrane proteins, much less how to measure their relative abundance. From a philosophical standpoint, we prefer to let cells tell us what sorts of broken protein variants are degraded by their QC systems, then do our best to surmise what this tells us about the relevant properties of cellular DSEs.

Despite this fundamental caveat, we believe that the chosen mutations and our interpretation of their relevant conformational effects are reasonably well-informed by current modeling tools and by prevailing knowledge on the physicochemical drivers of membrane protein folding and misfolding. Specifically, the mechanistic constraints of translocon-mediated membrane integration provide an understanding of the types of mutations that are likely to disrupt cotranslational folding. Though we are still learning about the protein complexes that mediate membrane translocation (https://doi.org/10.1038/s41586-022-05336-2), it is known that this underlying process is fundamentally driven by the membrane depth-dependent amino acid transfer free energies (https://doi.org/10.1146/annurev.biophys.37.032807.125904). This energetic consideration suggests introducing polar side chains near the center of a nascent TMDs should almost invariably reduce the efficiency of topogenesis. To confirm this in the context of TMD6 specifically, we utilized a well-established biochemical reporter system to confirm that V276T attenuates its translocon-mediated membrane integration (Fig. S1)- at least in the context of a chimeric protein. We also constructed a glycosylation-based topology reporter for full-length GnRHR, but ultimately found its' in vitro expression to be insufficient to detect changes in the nascent topological ensemble. In contrast to V276T, the W107A mutation is predicted to preserve the native topological energetics of GnRHR due to its position within a soluble loop region. W107A is also unlike V276T in that it clearly disrupts tertiary interactions that stabilize the native structure. This mutation should preclude the formation of a structurally conserved hydrogen bonding network that has been observed in the context of at least 25 native GPCR structures (https://doi.org/10.7554/eLife.5489). However, without a relevant folding assay, the extent to which this network stabilizes the native GnRHR

fold in cellular membranes remains unclear. Overall, we admit that these limitations have prevented us from measuring how much V276T alters the efficiency of GnRHR topogenesis, how much the W107A destabilizes the native fold, or vice versa. Nevertheless, given these design principles and the fact that both reduce the plasma membrane expression of GnRHR, as expected, we are highly confident that the structural defects generated by these mutations do, in fact, promote misfolding in their own ways. We also concede that the degree to which these mutagenic perturbations are indeed selective for specific folding processes is somewhat uncertain. However, it seems exceedingly unlikely that these mutations should disrupt topogenesis and/ or the folding of the native topomer to the exact same extent. From our perspective, this is the most important consideration with respect to the validity of the conclusions we have made in this manuscript.

> *Furthermore, plasma membrane expression has been used as a proxy for incorrect membrane protein folding, but this not necessarily be the case, as even correctly folded membrane proteins may not be trafficked correctly, at least, under heterologous expression conditions. In addition, mutations can affect trafficking and potential post-translational modifications, like glycosylation.*

While the reviewer is correct that the sorting of folded proteins within the secretory pathway is generally inefficient, it is also true that the maturation of nascent proteins within the ER generally bottlenecks the plasma membrane expression of most α-helical membrane proteins. Our group and several others have demonstrated that the efficiency of ER export generally appears to scale with the propensity of membrane proteins to achieve their correct topology and/ or to achieve their native fold (see https://doi.org/10.1021/jacs.5b03743 and https://doi.org/10.1021/jacs.8b08243). Notably, these investigations all involved proteins that contain native glycosylation and various other post-translational modification sites. While we cannot rule out that certain specific combinations of mutations may alter expression through their perturbation of post-translational GnRHR modifications, we feel confident that the general trends we have observed across hundreds of variants predominantly reflect changes in folding and cellular QC. This interpretation is supported by the relationship between observed trends in variant expression and Rosetta-based stability calculations, which we identified using unbiased unsupervised machine learning approaches (compare Figs. 6B & 6D).

> ***Reviewer #2 (Public Review):***
>
> *Summary:*
>
> *In this paper, Chamness and colleagues make a pioneering effort to map epistatic interactions among mutations in a membrane protein. They introduce thousands of mutations to the mouse GnRH Receptor (GnRHR), either under wild-type background or two mutant backgrounds, representing mutations that destabilize GnRHR by distinct mechanisms. The first mutant background is W107A, destabilizing the tertiary fold, and the second, V276T, perturbing the efficiency of cotranslational insertion of TM6 to the membrane, which is essential for proper folding. They then measure the surface expression of these three mutant libraries, using it as a proxy for protein stability, since misfolded proteins do not typically make it to the plasma membrane. The resulting dataset is then used to shed light on how diverse mutations interact epistatically with the two genetic background mutations. Their main conclusion is that epistatic interactions vary depending on the degree of destabilization and the mechanism through which they perturb the protein. The mutation V276T forms primarily negative (aggravating) epistatic interactions with many mutations, as is common to destabilizing mutations in soluble proteins. Surprisingly, W107A forms many positive (alleviating) epistatic interactions with other mutations. They further show that the locations of secondary mutations correlate with the types of epistatic interactions they form with the above two mutants.*

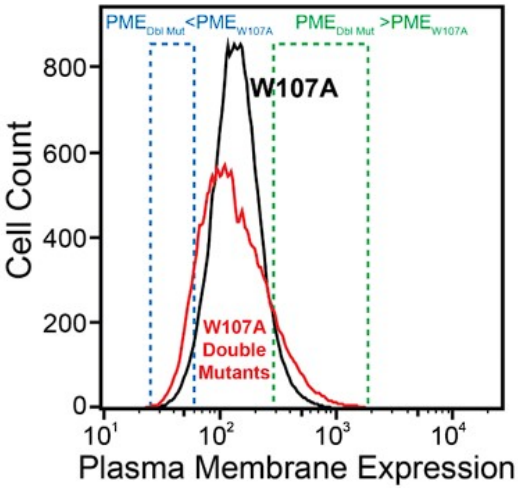
> *Strengths:*
>
> *Such a high throughput study for epistasis in membrane proteins is pioneering, and the results are indeed illuminating. Examples of interesting findings are that: (1) No single mutation can dramatically rescue the destabilization introduced by W107A. (2) Epistasis with a secondary mutation is strongly influenced by the degree of destabilization introduced by the primary mutation. (3) Misfolding caused by mis-insertion tends to be aggravated by further mutations. The discussion of how protein folding energetics affects epistasis (Fig. 7) makes a lot of sense and lays out an interesting biophysical framework for the findings.*
>
> *Weaknesses:*
>
> *The major weakness comes from the potential limitations in the measurements of surface expression of severely misfolded mutants. This point is discussed quite fairly in the paper, in statements like "the W107A variant already exhibits marginal surface immunostaining" and many others. It seems that only about 5% of the W107A makes it to the plasma membrane compared to wild-type (Figures 2 and 3). This might be a low starting point from which to accurately measure the effects of secondary mutations.*

The reviewer raises an excellent point that we considered at length during the analysis of these data and the preparation of the manuscript. Though we remain confident in the integrity of these measurements and the corresponding analyses, we now realize this aspect of the data merits further discussion and documentation in our forthcoming revision, in which we will outline the following specific lines of reasoning.

> *Still, the authors claim that measurements of W107A double mutants "still contain cellular subpopulations with surface immunostaining intensities that are well above or below that of the W107A single mutant, which suggests that this fluorescence signal is sensitive enough to detect subtle differences in the PME of these variants". I was not entirely convinced that this was true.*

We made this statement based on the simple observation that the surface immunostaining intensities across the population of recombinant cells expressing the library of W107A double mutants was consistently broader than that of recombinant cells expressing W107A GnRHR alone (see Author response image 1 for reference). Given that the recombinant cellular library represents a mix of cells expressing ~1600 individual variants that are each present at low abundance, the pronounced tails within this distribution presumably represent the composite staining of many small cellular subpopulations that express collections of variants that deviate from the expression of W107A to an extent that is significant enough to be visible on a log intensity plot.

Author response image 1.

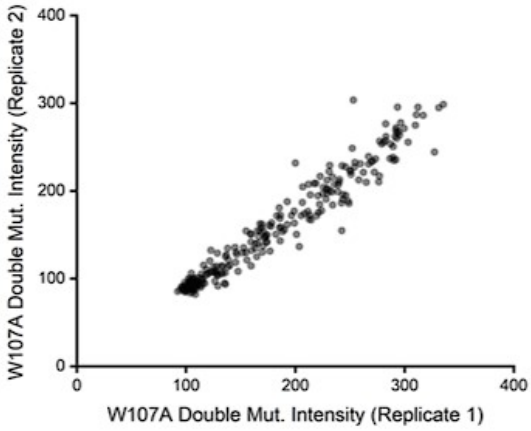


> *Firstly, I think it would be important to test how much noise these measurements have and how much surface immunostaining the W107A mutant displays above the background of cells that do not express the protein at all.*

For reference, the average surface immunostaining intensity of HEK293T cells transiently expressing W107A GnRHR was 2.2-fold higher than that of the IRES-eGFP negative, untransfected cells within the same sample- the WT immunostaining intensity was 9.5-fold over background by comparison. Similarly, recombinant HEK293T cells expressing the W107A double mutant library had an average surface immunostaining intensity that was 2.6-fold over background across the two DMS trials. Thus, while the surface immunostaining of this variant is certainly diminished, we were still able to reliably detect W107A at the plasma membrane even under distinct expression regimes. We will include these and other signal-to-noise metrics for each experiment in a new table in the revised version of this manuscript.

Beyond considerations related to intensity, we also previously noticed the relative intensity values for W107A double mutants exhibited considerable precision across our two biological replicates. If signal were too poor to detect changes in variant expression, we would have expected a plot of the intensity values across these two replicates to form a scatter. Instead, we found DMS intensity values for individual variants to be highly correlated from one replicate to the next (Pearson's R= 0.97, see Author response image 2 for reference). This observation empirically demonstrates that this assay consistently differentiated between variants that exhibit slightly enhanced immunostaining from those that have even lower immunostaining than W107A GnRHR.

Author response image 2.



> *But more importantly, it is not clear if under this regimen surface expression still reports on stability/protein fitness. It is unknown if the W107A retains any function or folding at all. For example, it is possible that the low amount of surface protein represents misfolded receptors that escaped the ER quality control.*

While we believe that such questions are outside the scope of this work, we certainly agree that it is entirely possible that some of these variants bypass QC without achieving their native fold. This topic is quite interesting to us but is quite challenging to assess in the context of GPCRs, which have complex fitness landscapes that involve their propensity to distinguish between different ligands, engage specific components associated with divergent downstream signaling pathways, and navigate between endocytic recycling/ degradation pathways following activation. In light of the inherent complexity of GPCR function, we humbly suggest our choice of a relatively simple property of an otherwise complex protein may be viewed as a virtue rather than a shortcoming. Protein fitness is typically cast as the product of abundance and activity. Rather than measuring an oversimplified, composite fitness metric, we focused on one variable (plasma membrane expression) and its dominant effector (folding). We believe restraining the scope in this manner was key for the elucidation of clear mechanistic insights.

> *The differential clustering of epistatic mutations (Fig. 6) provides some interesting insights as to the rules that dictate epistasis, but these too are dominated by the magnitude of destabilization caused by one of the mutations. In this case, the secondary mutations that had the most interesting epistasis were exceedingly destabilizing. With this in mind, it is hard to interpret the results that emerge regarding the epistatic interactions of W107A. Furthermore, the most significant positive epistasis is observed when W107A is combined with additional mutations that almost completely abolish surface expression. It is likely that either mutation destabilizes the protein beyond repair. Therefore, what we can learn from the fact that such mutations have positive epistasis is not clear to me. Based on this, I am not sure that another mutation that disrupts the tertiary folding more mildly would not yield different results. With that said, I believe that the results regarding the epistasis of V276T with other mutations are strong and very interesting on their own.*

We agree with the reviewer. In light of our results we believe it is virtually certain that the secondary mutations characterized herein would be likely to form distinct epistatic interactions with mutations that are only mildly destabilizing. Indeed, this insight reflects one

of the key takeaway messages from this work- stability-mediated epistasis is difficult to generalize because it should depend on the extent to which each mutation changes the stability (ΔΔG) as well as initial stability of the WT/ reference sequence (ΔG, see Figure 7). Frankly, we are not so sure we would have pieced this together as clearly had we not had the fortune (or misfortune?) of including such a destructive mutation like W107A as a point of reference.

> *Additionally, the study draws general conclusions from the characterization of only two mutations, W107A and V276T. At this point, it is hard to know if other mutations that perturb insertion or tertiary folding would behave similarly. This should be emphasized in the text.*

We agree and will be sure to emphasize this point in the revised manuscript.

> *Some statistical aspects of the study could be improved:*
>
> 1. *It would be nice to see the level of reproducibility of the biological replicates in a plot, such as scatter or similar, with correlation values that give a sense of the noise level of the measurements. This should be done before filtering out the inconsistent data.*

We thank the reviewer for this suggestion and will include scatters for each genetic background like the one shown above in the supplement of the revised version of the manuscript.

> 1. *The statements "Variants bearing mutations within the C- terminal region (ICL3-TMD6-ECL3-TMD7) fare consistently worse in the V276T background relative to WT (Fig. 4 B & E)." and "In contrast, mutations that are 210 better tolerated in the context of W107A mGnRHR are located 211 throughout the structure but are particularly abundant among residues 212 in the middle of the primary structure that form TMD4, ICL2, and ECL2 213 (Fig. 4 C & F)." are both hard to judge. Inspecting Figures 4B and C does not immediately show these trends, and importantly, a solid statistical test is missing here. In Figures 4E and F the locations of the different loops and TMs are not indicated on the structure, making these statements hard to judge.*

We apologize for this oversight and thank the reviewer for pointing this out. We will include additional statistical tests to reinforce these conclusions in the revised version of the manuscript.

> 1. *The following statement lacks a statistical test: "Notably, these 98 variants are enriched with TMD variants (65% TMD) relative to the overall set of 251 variants (45% TMD)." Is this enrichment significant? Further in the same paragraph, the claim that "In contrast to the sparse epistasis that is generally observed between mutations within soluble proteins, these findings suggest a relatively large proportion of random mutations form epistatic interactions in the context of unstable mGnRHR variants". Needs to be backed by relevant data and statistics, or at least a reference.*

We will include additional statistical tests for this in the revised manuscript and will ensure the language we use is consistent with the strength of the indicated statistical enrichment.