Super-compensator alleles rescue diverse deleterious substitutions and buffer the effects of mutations

Wenhong Jiang[1 † *], Hengchuang Yin[1 †], Saad S. Khan[1,2 †] and Yiparemu[3]

1 School of Life Science and Peking-Tsinghua Center for the Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

2 Departments of Bioengineering, Biology, Chemistry and Chemical Biology, Single Cell Proteomics Center, and Barnett Institute, Northeastern University, Boston, MA 02115, US

3 Henan University School of Stomatology, Kaifeng, China

† These authors have contributed equally to this work and share first authorship.

* Corresponding author: Wenhong Jiang: jiangwh@pku.edu.cn

Classification

Major: Biological Sciences.
Minor: Genetics

Keywords

compensatory mutations, deep mutational scanning, yeast, evolution, fitness

Abstract

Understanding loss-of-function mutations and their potential compensatory mechanisms is crucial for elucidating the complexities of gene expression. Here, we investigate whether additional mutations can rescue non-functional alleles and their evolutionary significance. To achieve this, we conducted an extensive analysis of 372,867 amino acid sequence variants of the imidazole glycerol-phosphate dehydratase (IGPD, PDB: 6ezm) protein, which is over 1000 times greater than previously examined. We discovered that one or more additional substitutions located in loop and interface regions could easily restore function for each of the 10,019 of 31,701 (31.6%) non-functional missense allele. We further identified a small set of super-compensatory substitutions that can increase the function of many

diverse alleles and flatten the fitness landscape, that is, buffer the fitness effects of both beneficial and deleterious substitutions. We also found that the protein structures of rescuable variants are more stable than those of non-rescuable variants. This suggests that restoring the stability of protein structures is a key part of compensation. These findings offer new insights into the understanding of protein function loss and its adaptive compensatory evolution.

Introduction

The majority of random amino acid mutations destabilize proteins and are slightly deleterious to their function (Bershtein et al. 2006, Zeldovich et al. 2007, Faure et al. 2015). Nevertheless, such mutations can accumulate within populations through genetic, molecular or ecological mechanisms. Molecular mechanisms such as epistasis can mask and alter the fitness defect of mutations within specific genetic backgrounds (Morton et al. 1956, Palmer et al. 2015). Many proteins exhibit mutational robustness, maintaining stability and function whilst accumulating mutations (Bloom et al. 2007, Lind et al. 2017). Population level dynamics enable the accumulation of deleterious mutations through genetic drift (Boeke et al. 1987, Fay et al. 2001, Szendro et al. 2013) or hitchhiking of deleterious mutations along beneficial ones (Lang et al. 2013). At the organismal level, allelic diversity can buffer deleterious effects owing to redundancies in expression when only a subset of the alleles accrue mutations (Morton et al. 1956, Fay et al. 2001). Finally, strong selective pressures such as drug treatments can create contexts where otherwise deleterious mutations confer an adaptive advantage promoting their retention within populations (Maisnier-Patin et al. 2004).

When the deleterious mutation is a single nucleotide substitution, there's a chance for a reverse mutation to restore the original phenotype. However, in cases involving multiple mutations, non-reversible mutations (eg, indels), or when environmental selection pressures favor the initial deleterious mutation, additional compensatory mutations offer another route to regain fitness. These additional mutations can bring the organism's fitness back to wild-type or nearly wild-type levels **(Figure 1A)**. This raises an important evolutionary question: once deleterious mutations reduce or inactivate a gene's function, what is the likelihood of a return to functionality, and what mechanisms might enable this reversion? Indeed, compensatory mutations as a route to restore the function of mutated genes have been widely reported (Maisnier-Patin et al. 2004, Bloom et al. 2010, Kimura 1985, Kondrashov et al. 2002, Tomala et al. 2019, Storz 2018). More notably, prior work has shown that compensatory mutations occur at a significantly higher rate than direct reversals (Poon et al. 2005). Furthermore, evolutionary experiments have shown that compensatory mutations do not simply occur after deleterious ones, to mask their effect. Instead, they often arise as neutral mutations that enable otherwise deleterious ones, acting as stepping stones that open evolutionary pathways inaccessible without them (Ortlund et al. 2007, Zheng et al.

2019). More so, in pathogens both theoretical and experimental evidence demonstrate that compensatory mutations restore organismal fitness, especially following population bottlenecks induced by antibiotic treatment (Meftahi et al. 2016, Comas et al. 2011, Handel et al. 2006, Liu et al. 2018).

The molecular mechanisms of intragenic compensatory mutations can be broadly divided into those which act globally and those with a local effect. Global compensatory mutations counteract the structural destabilization induced by deleterious mutations. As such, these mutations have fewer constraints relating to the distance between mutations, so long as they occur within the same protein domain. This flexibility allows for generalized compensatory substitutions capable of rescuing a wide range of deleterious mutations (Tomala et al. 2019). In contrast, local compensatory mutations involve specific physical interactions between residues, such as those within an active site or a single α-helix. These mutations exert highly localized and specific effects, offering less generalizable compensatory capabilities (Ivankov et al. 2014). As such, in the absence of the primary mutation, these mutations can themselves introduce destabilizing and deleterious effects (Poon et al. 2005).

The rescue of deleterious genotypes in small scale, controlled experiments has been previously studied at the level of proteins (Maisnier-Patin et al. 2004, Bloom et al. 2010, Hanson et al. 1993, Mateu et al. 1999), mRNA's (Stephan et al. 1993, Kirby et al. 1995), and tRNA's (Domingo et al. 2018, Poon et al. 2006). While the studies did demonstrate compensatory mutations, they also highlighted certain experiments where fitness was only restored through reversion rather than via compensatory mutations (Filteau et al. 2015, Jakubowska et al. 2009). Despite their limited scale, the studies did report a few general mechanisms relating to rescue. For example, deleterious mutations are rescued with higher frequency when the mutated residue is surface-exposed and interacts with fewer other residues (Mateu et al. 1999, Ferrer-Costa et al. 2007). Similarly, when single mutations occur at conserved sites, they tend to thermally destabilize a protein and are generally only rescued by additional mutations occurring in conserved sites (Tomala et al. 2019). However, the limited scope of these studies, which examined fewer than 100 genotypes, present cross study observational inconsistencies and thus constrains the elucidation of broader impact and mechanisms of compensatory mutations (Tomala et al. 2019, Poon et al. 2006, Ferrer-Costa et al. 2007). As such, it remains challenging to generalize mechanisms of compensation without reliably identifying biochemical or structural features that are predictive of rescuability via compensatory mutation.

To address this gap, we used a Deep Mutational Scanning (DMS) dataset of ~400,000 amino acid genotype variants of the yeast HIS3 gene. The gene codes for the conditionally essential enzyme

imidazoleglycerol-phosphate dehydratase (IGPD/His3p) that is required for histidine biosynthesis (Pokusaeva et al. 2019). The amino acid variants span most of proteins sequence and 30,000 variants render the yeast unable to grow in media lacking histidine. This dataset, which is over 1000 times larger than those of previous studies, allowed us to determine the structural and biochemical rules which govern compensatory evolution, and determine molecular mechanisms by which additional substitutions can rescue unfit genotypes.

Results

A DMS to identify the structural and physicochemical rules of intragenic compensation.

To quantitatively explore the complete spectrum of compensatory mutations across an entire protein, we analyzed an evolutionary DMS library of HIS3 gene from the *S. cerevisiae***.** In this library, to generate locally dense and combinatorically complex groups of sequences, the protein was divided into twelve independent segments of 20-30 amino acids each, and each segment was mutagenized and fitness measured experimentally in a pooled competition assay (Pokusaeva et al. 2019). We scaled the normalized relative fitness of each genotype to zero for 95% of genotypes with premature stop codons and to one for WT genotypes (**Materials and Methods**). To obtain a set of reliable data, we discarded experimental data in which the correlation between the two experimental biological replicates was less than 0.85 and discarded genotypes with fewer than three synonymous variants. This makes a library with 372,867 variant genotypes of amino acids. Most of them (74.3%) are combinations of amino acid that are found in 21 different types of living yeast. **(Fig. 1B, S1A-B).** On average, the genotypes have 7 amino acid substitutions relative to *S. cerevisiae* reference, with each amino acid genotype coded for by an average of 9 nucleotide synonymous variants, which serve as within-experiment biological replicates. Of the 372,867 genotypes, 31,701 (8.5%) exhibited a relative fitness of zero, which provides the principles underlying intragenic compensation.

Visually, it's evident that many unfit genotypes can be rescued by additional substitutions **(Fig. 1B)**. For instance, a genotype shown in blue in Fig. 1B has four substitutions with a fitness of 0. However, of the 23 genotypes in our library that contain these same four substitutions plus one or more additional substitutions, nine have a fitness significantly greater than zero and rescue the unfit genotype **(Fig. 1C)**. To determine how often one unfit genotypes can be rescued, for each unfit genotype (fitness = 0), we identified all genotypes with the same set of substitutions plus one or more additional substitutions. We defined the rescuability of each unfit genotype as the sum of the fitness difference between the unfit genotype and all corresponding rescuing genotypes, divided by the total number of genotypes with

additional substitutions **(Fig. 1C, see methods)**. Such number is limited to minimum of ten to improve the reliability.

A major source of error in this analysis is 'false-unfit' – calling a genotype unfit when it's actually fit, and these false-unfit genotypes will have very high rescuability. To estimate such error, we analyzed nonsense genotypes with stop codons, assuming that additional substitutions won't restore function since they do not remove the stop codon. From the nonsense genotypes, we calculate a False Discovery Rate (FDR) of 5% at rescuability >= 0.02, and an FDR of 1% at rescuability >= 0.04. Far fewer nonsense genotypes can be rescued, and those that can, are much less rescuable compared to missense genotypes **(Fig. 1E)**. As the nonsense amino acid genotypes have few synonymous variants, we included nonsense genotypes with less than three nucleotide synonymous variants, and thus the fitness values have higher uncertainty. Thus, these are conservative estimates and the real FDRs are likely quite a bit lower. Considering the FDR from nonsense genotypes, we defined one unfit genotype that can be rescued when the rescuability is at least 0.1. Using a threshold of rescuability is >= 0.1, 11.8% (3750) of unfit genotypes can be rescued. This is the subset that was used for all further analysis, though using a threshold of 0.2 or 0.04 does not substantially affect the results.

Our conclusions remain consistent using various methods to calculate fitness and applying different thresholds and filters **(Fig. S2A-F)**. Moreover, the rescuability and the fraction of rescuable genotypes are hardly affected by the number of genotypes at each position or the total number of genotypes with additional substitutions (**Fig. S1F**).

Characterization of features that influence whether an unfit genotype can be rescued

To understand the structural and physicochemical features determinants of intragenic compensation, we took advantage of the fact that our library contains thousands of non-functional genotypes with substitutions spread across almost the entire protein **(Fig. S1E),** covering a wider range of protein structural features and biochemical properties, allowing us to determine the general principles regarding intragenic compensation.

To calculate rescuability for each single amino acid position in His3p, for each position, we calculated the average rescuability of all unfit genotypes that contain a substitution at this position. We found large differences in average rescuability across the protein, with disordered loops, especially the surface loop from position 30 to 47 and the C-terminal loop that coordinates the active site of the enzyme (204 - 220) (38), having extremely high rescuability**(Fig. 2A)**, while those in β-strands are more difficult to rescue **(Fig. 2B) (Fig. S3A-B)**. Similar findings were observed for the IgG-binding domain of protein G(GB1) (Wu et al. 2016), as shown in our analysis results following the same pipeline (**Fig. S3C**).

When map those substitutions to protein 3D structure, we found that solvent-exposed positions on the surface of the protein as well as those less located in the protein's core (which are closer to other

residues) tend to be more rescuable (**Fig. 2C).** Furthermore, less evolutionarily conserved positions are more easily rescued **(Fig. 2C)**, consistent with the previous work (Ferrer-Costa et al. 2007).

One striking observation from having unfit genotypes along nearly the entire length of the protein was that the rescuability of the C-terminal half was higher than that of the N-terminal half **(Fig. 2A)**. Given that IGPD is a 220 amino acid protein that forms a homo-24-mer, and each subunit is a tandem duplication of about 100 amino acids, it was surprising to find that the structurally similar C-terminal half had significantly higher rescuability. This higher rescuability may be attributed to the C-terminal half forming the majority of inter-subunit contacts within the homo-24-mer **(Fig. 2A, black dots)**, given that unfit genotypes with substitutions in the interface regions are more easily rescued (**Fig. 2D)**. This suggests that while mutations can easily disrupt subunit-subunit or protein-protein interactions, restoring functional interactions through additional mutations is also relatively straightforward.

Non-rescuable genotypes are more severely affected by stability impairment, compared with genotypes that can be rescued

To understand the mechanism of intragenic compensation from the viewpoint of stability, we evaluated the change in free energy related to both structure and substrate binding across a representative subset (10619) of all genotypes considering the limited computation resource. We found that rescuable genotypes are more stable than those which can't be rescued, independent of the number of substitutions (**Fig. 3A, Fig. S5)**. While rescuing genotypes are more stable than the unfit rescuable genotype, suggesting that compensatory mutations can act via stabilization, this was truer for slightly destabilized genotypes ($\triangle\triangle G < 40$). However, the predicted stability was rarely fully restored by compensatory substitutions (**Fig. 3B**).

For the substitutions in the interface regions, the stability of monomer or homo-24mer can be affected, suggesting the possibility of polymorphic effects for those substitutions (Tokuriki et al. 2009, Stein et al. 2019, Casadio et al. 2011). We calculated the change of binding energy for genotypes that have more than 50% substitution in the interface regions, as we expected, the rescuable genotypes have high binding affinity stability than non-rescuable genotypes, and rescuing genotypes can increase the stability of homomer binding affinity (**Fig. 3C-D, Fig. S5B**). In conclusion, not only does intragenic compensatory substitution restore protein stability, but genotypes that can be rescued tend to be more stable, and this is true for both monomer stability and change in free energy upon homomerization.

Changes in amino acid physicochemical properties can predict fitness and rescuability at specific positions

The physicochemical properties of amino acids (Kawashima et al. 2008) have been widely used to predict the stability effects of mutations successfully and even greatly increase the accuracy of a single mutation on protein stability prediction (Teng et al. 2010, Huang et al. 2007, Liu et al. 2012, Fang et al.

2019). Combined with the fact that compensation for protein stability can mostly explain this data, we hypothesized that the compensatory substitutions should also bring the physicochemical properties of unfit variants back to the WT state. To test this, we used the AAIndex database of physicochemical properties for each amino acid and calculated all 553 AAIndex scores for each genotype, as the absolute difference between the mutant's score and the wild-type score (see methods, AAindex value distance to WT). This is a measure of the distance of each genotype from the wild-type, not in mutational space, but in physicochemical space. For each group of genotypes with the same substitutions, we found the relationship between fitness and the distance to the wild-type in physicochemical space **(Fig. 4A)**. We found that a single AAIndex feature was often highly correlated with fitness at one set of mutated positions, while completely uncorrelated with fitness at other positions **(Fig. 4A)**. Intuitively, this makes sense; different physicochemical properties are important in different structural and functional regions of the protein (Sengupta et al. 2012, Narayanan et al. 2013, Jaspard et al. 2014). For example, the types of interactions between residues vary across different regions of proteins (Alauddin et al. 2015, Jelesarov et al. 2009), and only specific physicochemical properties contributions to such interaction (Sengupta et al. 2012, Lazar et al. 2018).

We first found the correlation coefficient matrix for fitness and the 553 AAindex values distances to WT for each set of genotypes with substitutions at the same positions across the whole protein. This helped us see how physicochemical properties relate to fitness in different positions and across the whole protein. We clustered this matrix based on the similarity of the correlation coefficients of different physicochemical properties **(Fig. 4C)**. This reveals two trends. First, most of the physicochemical properties (447/553 of them) are more linked to fitness in the structured non-loop parts of the protein than they are in the loop parts. **(Fig. 4C-D)**. Second, most physicochemical properties are predictive of fitness at only a small fraction of positions in the protein **(Fig. 4C)**. For example, KARS160109, which is related to the size of the sidechain, is highly predictive of fitness for genotypes with mutations at positions 106,108,110,111,112,131 and 135, 71.4% of which are in loop regions **(Fig. 4A-B)**, while CRAJ730102, related to the frequency of each amino acid occurs in beta-sheets, is predictive for genotypes with substitutions at positions 7,8,9,19,20,21,22 and 23 (**Fig. 4A-B**). There is no correlation between AAindex values distance to WT and fitness after the same analysis based on the same data with the fitness of genotypes randomly shuffled (**Fig. S6**).

The high correlations between amino acid physicochemical properties and fitness of genotype suggest the underlying mechanism of compensation. For example, if the change in the value of the feature CRAJ730102 spread over mentioned above nine amino acid positions is causal for fitness, then, starting from a single genotype, it can compensate for genotypes from a subset (genotypes that have fewer substitutions relative to the single starting genotype) by additional substitutions which move the value closer to that of the wild type, while additional substitutions that move the value distant should decrease fitness. And we call this phenomenon as the feature can predict compensation **(Fig. 5A)**. Similarly, we

can look at the superset – genotypes that have one or more substitutions relative to the single starting genotype.

To test this, we did the following analysis; firstly, for each correlation between fitness and feature value distance to WT in the original genotypes set, we found out the total number of genotypes in the subset as (N1+N2), where N1 represents the number of genotypes for which one single feature predicts compensation, N2 doesn't. After that, we calculated the ratio of N1/(N1+N2) **(Fig. 5A)**. Indeed, we find that the stronger the correlation between fitness and feature value distance to WT in the original set, the more genotypes can predict compensation by a single feature in the subset **(Fig. 5B),** and it is also the same for each physicochemical feature in the superset. In general, when a physicochemical feature can predict fitness in some specific regions of His3p, this feature can also predict the ability of additional substitutions to rescue fitness. It also shows that feature values at individual positions are relatively unimportant. Instead, the importance is distributed across multiple residues, which means deleterious changes at one position can be compensated by rescuing changes at another position that restore the total summed feature value to that of the WT.

Super compensatory substitutions can increase the fitness of diverse alleles, and super compensators can flatten the fitness landscape by reducing the fitness effects of substitutions

We next sought to determine which types of single amino acid substitutions can rescue genotypes from different genetic backgrounds. There are two broad possibilities. The first is that rescue is highly genotype-specific, that is to say; individual substitutions will rescue only a limited number of unfit genotypes or only genotypes with specific substitutions. The previous discussion suggests a highly specific relationship between fitness and amino acid physiochemical properties, which supports this model. The second is that individual substitutions would rescue a wide variety of genotypes.

We, therefore, calculated how often every substitution from wild-type to an alternative allele can restore fitness in the presence of each non-WT amino acid state (**Fig. 6A,** for example). The relationship between each amino acid state and substitution from WT can be quantified as a directed graph in which a directed edge from A to B represents the ability of substitution B to increase fitness in the presence of amino acid state A, and the weight of that edge is the fraction of genotypes in which substitution B increases the fitness at least 0.5 in the presence of amino acid state A (**Fig. 6A-B**). Therefore, the sum of the weights of incoming edges is a measure of the diversity of genetic backgrounds from which a given substitution can restore fitness. We note that this is different from the distribution of fitness effects for each substitution, as it downweighs specificity. If a given substitution can only rescue a limited diversity of genetic backgrounds, there will be fewer incoming edges. Similarly, the graph method normalizes for the library's diversity, which means that individual amino acid state varies in abundance, and more common amino acid states will count for more when calculating the distribution of fitness

effects. We calculated the compensatory ability for every substitution from wild-type by summing all the weights of edges pointing to the same substitution. We found that a limited number of substitutions can increase fitness in a wide variety of backgrounds with high compensatory ability (**Fig. 6C-D, Fig. S8**). In those super compensatory substitutions, we refer to these non-WT amino acid states as super compensators.

One obvious hypothesis is that super compensatory substitutions cause a significant increase in fitness across a wide array of genetic backgrounds. To investigate this, we used a previously developed neural network model (Pokusaeva et al. 2019) in which each amino acid state is given a weight that measures its contribution to an intermedia variable, called fitness potential. A neural network then maps fitness potential to fitness (**Fig. 6E-F**), following the sigmoid function. Thus, each amino acid state receives one weight, which we call the fitness impact score, and each genotype receives a value for fitness potential, which is simply the sum of the weights, and a predicted fitness. Substituted amino acid states that have large positive fitness impact scores than WT in the model, can increase fitness potential as well as fitness across most genetic backgrounds, thus explaining the existence of super-compensators **(Fig. 6F-G)**.

Following this, we explored how the existence of a super-compensator affects the distribution of fitness effects. We noticed that both deleterious and beneficial substitutions tended to be neutral when one wild-type state in another position was replaced with a super-compensator **(Fig. 7A-B)**. To quantify this globally, we compared the fitness effect of each deleterious or beneficial substitution when that substitution occurs in the wild-type state **(Fig. 8C**, y-axis**)** to when the same substitution appears in the super-compensator **(Fig. 8C**, x-axis**)**. Each point represents the calculation from four genotypes that differ at only two positions: one being the position of substitution, and the other either a wild-type amino acid state or a super-compensator **(Fig. 8A)**. When substitutions exhibit similar fitness effect in both WT and super compensator alleles, the points are evenly distributed around the diagonal **(Fig. 8E-F)**, and the skewness of the difference between the x-axis value and the y-axis value of each point is around 0**(Fig. 8I-J, L)**. However, for the super-compensator allele like 110D, the distribution appears almost vertical **(Fig. 8C-D)**, suggesting F.E. has been greatly impacted. To quantify the distribution differences, we then calculate the difference between y-axis and x-axis values for each point, separately for fitness-increasing and fitness-decreasing substitutions. Surprisingly, we observed that substitutions, whether increasing or decreasing fitness, exhibit weaker fitness effects and greater skewness when occurring in a super-compensator background **(Fig. 8G-H, 8K)**. For simplification, in the following analysis, we only considered the absolute F.E. for substitutions regardless they are beneficial or deleterious **(Fig. 8K-L)**.

To understand the cause, we return to the sigmoidal relation between fitness potential and fitness. Most genotypes, in our library, and nature, have high fitness, but their sequence varies. There is a plateau in which fitness potential can vary, but fitness remains roughly constant and high **(Fig. 8B)**. There is a

similar lack of sensitivity at zero fitness where if a genotype is already unfit, additional deleterious substitutions will not affect fitness **(Fig. 8B)**. Beneficial substitutions decrease the sensitivity by moving many genotypes to the right on the fitness potential plateau, resulting in positive skewness **(Fig. 8B-C)**. Deleterious substitutions move genotypes away from the plateau and increase the sensitivity, resulting in negative skewness **(Fig. 8B, D)**. When we used the fitness impact score to quantify the ability of one substituted amino acid state to affect the fitness, we found that if amino acid states have a larger fitness impact score than WT, the skewness increases (**Fig. 8F**). Moreover, the greater the compensatory ability of substituted amino acid states, the larger the skewness observed (**Fig. 8E**). Therefore, the reason that super-compensators buffer the effects of both deleterious and beneficial substitutions is the flattened fitness landscape on the fitness potential plateau.

Compared to studies that uncover the rugged natural features of the fitness landscape (Pokusaeva et al. 2019, Bak et al. 1992, Van Cleve et al. 2015, Kondrashov et al. 2015, de Visser et al. 2018), we found that super-compensators can flatten the fitness landscape by reducing the F.E. of substitutions. As illustrated in Figure 9, the super compensator 93V can flatten the local fitness landscape of variants with positions 66 and 86 substituted. It is worth mentioning that we found other super compensator-like alleles in another protein – avGFP (Rawson et al. 2018) random library based on a similar method that assigns a fitness impact score to each amino acid state using deep learning models. These super compensator-like alleles can also increase the fitness (fluorescence intensity) of diverse alleles compared to the wild-type amino acid state (**Fig. S10)**.

Coexistence of super compensators with deleterious amino acid states in extant species.

Super compensators can increase the fitness of diverse alleles, which suggests that super compensators can co-exist with unfit amino acid states, but the function of protein hardly be disturbed. This suggests that in different extant species where His3p plays as a normal and key enzyme, the diversity of His3p sequences may come from the combination of super compensators and other deleterious amino acid states. To test this idea, we analyzed 335 His3p sequences from extant species which show more than 50% identity with the His3p sequence of *S.cerevisiae*. And in each extant sequence, we scanned every pair of non-WT amino acid states compared to that of *S.cerevisiae*. We found that super compensators can co-exist with slightly deleterious amino acid states which have negative fitness impact scores relative to wild type amino acid states (**Fig. 9A, Fig. S11**). Further analysis demonstrated that within such pairs in extant species, the presence of a more deleterious amino acid state (with a lower fitness impact score relative to wild-type) is associated with a higher fitness impact score of the accompanying super compensator (**Fig. 9B**).

Super compensators can stabilize local interaction

The previous discussion revealed the importance of restoring protein stability in this data for compensation. Therefore, we screened the contribution of stabilizing protein structure for super compensators. We found that super compensators which we have selected can stabilize the local protein structure by different inter-residue interactions (**Fig. 10**). For example, the substitution A110D can help the formation of salt bridges with 112R, thus stabilizing the conformation of adjacent loop regions. In contrast, amino acid states which have low fitness impact scores relative to wild-type might strongly destroy the local structure, for example, the substitution I19D might destroy the β-sheet and spiral accumulation. We listed the structural effects of top7 super compensators which have the highest fitness impact score and 5 amino acid states which have the lowest fitness impact score (**Table S1**). Notably, most super compensators are not near the substrate or Mn^{2+} (< 8 **Å**) except for 189A, 137G, and 93V, but neither of them is within 4 **Å** distance of the substrate or Mn^{2+} （**Fig.S 12**）, indicating the effects to substrate or Mn^{2+} of super compensators are limited.

Super compensators increase mutational robustness to random mutations

The above results predict that super compensators could increase the robustness of the wild-type genotype to deleterious mutations. To test this hypothesis, we constructed two deep mutational scanning libraries. In each, we mutated two flanking positions to all other amino acids (NNK), and varied the third, central position to contain either the wild-type or a super compensator allele (**Fig. 11A**). We found that both super compensator alleles tested reduced the fitness impact of most mildly deleterious substitutions; 18 substitutions were deleterious only in the wild-type background, while only 3 substitutions were deleterious only in the super compensator background (**Fig. 11A, B**). For both super compensator alleles tested, deleterious substitutions had a smaller effect in the super compensator background than in the wild-type background (**Fig. 11C, D**) suggesting that super compensator alleles can flatten fitness landscapes, reducing the effects of many different types of deleterious substitutions.

Discussion

A unique advantage to analyze the deep mutational scanning library based on an evolutionary approach is that, since most sequence variation originates from extant species, 35% of genotypes with ten or more substitutions are still functional. This contrasts with random libraries, where the vast majority of genotypes with more than a few substitutions are unfit (Sarkisyan et al. 2016). Using this library, we conducted an in-depth study in two areas: first, identifying the characteristics of unfit but rescuable substitutions, and second, determining what rescuing substitutions look like.

In the first part, we found that the compensation phenomenon is very common. But the rescuability is greatly different for different deleterious substitutions. If deleterious substitutions are in less conserved sites, have fewer contact numbers with other residues, are more solvent-exposed, and in interface regions, these substitutions are easy to be compensated for. This can explain why Tomala et al (Tomala et al. 2019) only found a limited number of compensatory mutations for temperature-sensitive mutations at other sites because these temperature-sensitive mutations are located at folded and conserved sites. And in another independent experiment, these temperature-sensitive mutations can never be compensated (Jakubowska et al. 2009). And consistent with Tomala et al results, we found that deleterious substitutions which cause less structural stability can be rescued, and rescuing substitutions can offset such destabilizing effects, which suggests the destabilizing effect of substitution in folded and conserved positions is stronger than those in the loop and solvent-exposed positions, and this hypothesis has already been proved in the past ten years (Greene et al. 2001, Tokuriki et al. 2007). However, there are two points to note, one is that the data analyzed is built from extant species, and most of the amino acids in any genotype are from extant species. These amino acids kept by evolution are thought to have less deleterious effects on the function of the protein compared to random substitutions. And for those extremely conserved residues like XXXHHXXX (X: other amino acids) motif in both C-terminal and N-terminal regions, these residues never be substituted, thus we lacked related data for those extremely conserved positions. However, the results of the rescuability distribution indicate that substitutions in these highly conserved positions are unlikely to occur. Another point we address here is the potential for inter-segment compensation. We studied the inner compensation by splitting the His3p gene into 12 segments of similar length. Current technological limitations prevent the exploration of the entirety of genotypes, which are vast in number. On the one hand, the data in each segment makes the depth firstly rather than the width across the whole protein. On another hand, each segment can be considered as one individual gene. The interactions between segments to contribute to the function of the protein look like the interactions of genes to contribute to the overall organismal fitness. Inner-segment compensation of the whole protein can be considered as the intragenic compensation of the whole genome.

Next, we study if physicochemical features of amino acid can predict compensation. Surprisingly, we found that physicochemical features of amino acid alone can predict compensation as shown in Figure 4. However, in Figure 4C, there are some cells that have a strong positive correlation (correlation > 0.5, 6443/429681). As the data represented by these cells, the wild-type genotype should be unfit if predicted by these physicochemical features (**Fig. S7B)**. This led us to reconsider these physicochemical features should not be the causality of fitness. We screened these physicochemical features with positive correlation predicting fitness and found that other highly correlated physicochemical features can predict fitness even including the wild type genotype (**Fig. S7A-C**, for example) and also can predict

compensation(N1/(N1+N2) = 0.48 in the subset and N1/(N1+N2) = 0.48 in the superset). This observation hinted at the possibility that the initial positive correlations were perhaps a result of the AAindex matrix's redundancy rather than indicators of direct causation. Besides, these physicochemical features with positive correlations can't predict compensation.

Next, we study the features of compensatory substitutions. Surprisingly, we found that some substitutions can always compensate for many other alleles. But when we considered this phenomenon from the view of the change in protein stability, it is reasonable. Because these super compensatory substitutions can stabilize the local protein structure and thus offset the destabilizing effects of many other deleterious substitutions. The distance between compensatory substitutions and compensated substitutions revealed that the most compensation is not the restoration of the direct interaction between two residues (**Fig. S12**). Many studies have found that compensatory substitutions can offset the fitness cost of drug-resistant mutation (Meftahi et al. 2016, Comas et al. 2011, Handel et al. 2006, Liu et al. 2018). Combined with the study that cryptic mutations make more subsequent mutations kept in evolution (Zheng et al. 2019), we are convinced that super compensatory substitutions can also help the accumulation of subsequent deleterious mutations in a population, which makes the inaccessible evolution pathway accessible. And our experiment proved that. Our findings underscore the critical role of super compensators in shaping the fitness landscape—specifically, in flattening it—and in facilitating population evolution. The existence of super compensators can also explain the mutation robustness property of the protein. Both HaiWei H. Guo et al (Guo et al. 2004) and György Abrusán et al (Abrusán et al. 2016) found that the mutation robustness is different for different secondary structures of the protein, which is quite similar to our finding that deleterious substitution are easier to be compensated in loop regions. Super compensators give us the inspiration for novel protein design. If we can predict super compensators, we can design more novel proteins with function normal and potential functions. Based on our finding that super compensators can stabilize the local protein structure, we highly suggest that stabilizing substitutions search is important and necessary for novel protein design, which is in line with the prevailing protein design philosophy, and there are many works about the quick prediction of stabilizing mutations (Popov et al. 2018, Borgo et al. 2012, Zimmerman et al. 2017, Steipe et al. 1994).

However, although we tried to explain the mechanism of compensation by the stability of the structure and interface binding affinity, we still don't know the other possible effects of compensatory substitutions like the protein expression level, protein solubility, catalytic efficiency, and so on. Therefore, further work with higher resolution to study the compensation on a large data scale and high thought are needed.

Materials and Methods

Experiment design and data analysis to determine the fitness of each genotype

The experiments and analyses have already been fully described (Pokusaeva et al. 2019). A brief description follows. Firstly, each His3 gene variant that encodes IGPD with different kinds of extant amino acids was contracted into a plasmid and transformed into *S. cerevisiae* with the His3 gene deleted. Next, *S.cerevisiae* populations of which each yeast contains one kind of His3 gene variant were cultured together to compete so that substitutions in IGPD which are deleterious will grow slowly and decline over time. However, substitutions that are neutral or beneficial will grow fast and gradually prevail. Based on the competition, variants with missense substitutions that have similar or slower growth rates than variants with stop codons have the fitness of zero. And wild-type genotypes and most variants with one substitution have the fitness of 1. This experiment is conducted with two biological replicates and two technical replicates of each biological replicate.

Rescuability calculation

For each unfit genotype, we found all genotypes that have the same set of substitutions plus one or more additional substitutions (superset, the total number is N). For each genotype, we have the value of fitness (S) and uncertainty in our measurement of fitness (standard deviation, S.D.), with the actual value of fitness in the range S±1.96S.D. (95% confidence interval). We state that genotype A can compensate B when $\boldsymbol{S_A}$-1.96$\boldsymbol{S.D._A}$ >= $\mathbf{S}_B$ +1.96$\boldsymbol{S.D._B}$. The number of superset genotypes that can compensate is N1, and the fitness difference between the unfit genotype and each genotype in N1 is Δf. To calculate the rescuability of each unfit genotype we summed each Δf and divided it by the total number of superset genotypes, N.

$$Rescuability = \frac{\sum_1^{N1} \Delta f}{N}$$

Also, we used another method to make a statement of compensation. Genotype A has different synonymous nucleotide variants than G_A, and genotype B has different synonymous nucleotide G_B variants. For these two groups of nucleotide variants, we state that if frequencies of G_A are significantly bigger than G_B using the Kolmogorov-Smirnov test (KS test) with significance level α = 5%. Δf between rescuing and rescuable genotypes and rescuability can be calculated. The rescuability calculated by these two methods is highly correlated (**Fig. S2D**). In this manuscript, we used the former one.

We used different thresholds to calculate the average rescuability for each position in IGPD. The average rescuability for each position is the average rescuability of unfit genotypes with this position

substituted, and the unfit genotypes can be genotypes that have fitness equaling to zero, smaller than 0.5, or smaller than the fitness of wild type genotypes. And filtered the unfit genotypes according to the total number of genotypes that have additional substitutions, and the total number can be 0 or 10. Results are shown in FigureS2A-D.

WCN (weighted contact number) score, conservation score, RSA (Relative solvent accessibility) score calculation, and interface regions identification

We calculated the weighted contact number (WCN) described by Lin CP et al (Lin et al. 2008) for each residue in IGPD, and the formula can be written as below,

$$WCN_i = \sum_{i \neq j}^{N} \frac{1}{r2(ij)},$$

where r_{ij} represents the distance between the geometric center of the C_α carbon atoms of residues *i* and the geometric center of the C_α carbon atom of residues *j* in IGPD with N residues. We used the average WCN of 24 monomers for each position.

To calculate the conservation score, firstly we aligned all the amino acids sequences of one monomer of IGPD in extant 21 yeast species using T-coffee Expresso (Notredame et al. 2000) and for each position in IGPD, we calculated the entropy of amino acid composition as below,

$$H(X) = \sum_{x=0}^{20} P(x) log2P(x);$$

X represents one position in IGPD, P(x) represents the frequency of one kind of amino acid of 20 in position X in extant 21 yeast species. And conservation score is defined as follows,

$$C(X) = (1-H(x)/(1-gap_fr);$$

gap_fr represents the frequency of the gap in position X.

The RSA value for each residue in IGPD is recorded in the DSSP file (converted from the PDB file of IGPD (Rawson et al. 2018) by DSSP program (Touw et al. 2015, Kabsch et al. 1983)), and we normalized the RSA values by the surface area of the corresponding amino acid measured by Matthew et al(Tien et al. 2013). To identify the interface regions, we followed the standard PyMoL tutorial for interface residues identification with default parameters.

The calculation of energy change of protein conformation, substrate binding, and interface binding after substitutions

The energy change of interface binding followed the standard Rosetta tutorial for the calculation of the change in binding energy caused by mutations in the complex with default parameters.

Evaluation of physicochemical features can predict fitness and compensation

Each cell in Figure4C represents a spearman correlation between one physicochemical feature distance to WT of genotypes with the same position substituted (sample size >= 82 with the alpha of 0.05, the power of 0.80, and an effect size of 0.3 by using G*Power 3.1(Faul et al. 2009)) and the corresponding fitness. If the p-values of the spearman correlation are less than 0.05 when conducting the test that if the correlation coefficient equals zero, we set the correlation to zero to filter the unconvincing correlations.

Super compensators identification and buffering ability(skewness) calculation

For discussion, we abbreviate position *i* as *pos_i* here. For each substitution (S_{pos_i}) from the wild-type allele (WT_{pos_i}) and each non-wild-type allele N_{pos_j}, S_{pos_i} can compensate N_{pos_j} when a genotype (G_1) with S_{pos_i}+N_{pos_j} has higher fitness than a genotype (G_2) with WT_{pos_i}+N_{pos_j} and the two genotypes differ only at position *i*. We calculated the fitness difference $\triangle f$= fitness(G_2) – fitness(G_1), and the fraction of genotype pairs in which $\triangle f>0.5$. This fraction is stated as the compensatory ability of substitutions from WT_{pos_i} to S_{pos_i} to N_{pos_j}. Such a relationship can be quantified as a directed graph in which a directed edge from S_{pos_i} to N_{pos_j} represents the fraction of genotype pairs in which $\triangle f>0.5$. The graph visualization in Figure 6B and Figure S8A is achieved by *python* module NetworkX(76).

To determine the effect of super-compensators on the fitness-effect size of substitutions, we only considered the substitutions which cause the fitness change (fitness effect ≠ 0). The fitness effect of a given substitution is the difference in fitness between a pair of genotypes that differ only at that substitution (G_1=00000 -> G_2=00100). We calculate the fitness effects for all pairs of genetic backgrounds that differ only by one other position (G_1=00000 -> G_2=00100 & G_3=00001 -> G_4=00101), with each genotype G_1, G_2, G_3, and G_4 having fitness f_1,f_2,f_3, and f_4. $\Delta f_1 = f_2-f_1$ and $\Delta f_2 = f_4-f_3$. If the absolute value of $\Delta f2$ tends to be smaller than $\Delta f1$, then the single allele that is different between G_1 & G_3 (and G_2 & G_4) buffers the effects of mutations. We calculated the skewness of $\Delta f1$ - $\Delta f2$ to quantify the buffering ability of this single allele.

Deep mutational scanning library to test the ability of super compensators to confer mutational robustness

To experimentally validate our predictions, we selected two S.C. alleles with distinct potential structural mechanisms of rescue: A110D and S189A. The sites for random mutagenesis were selected based on

each allele's compensatory ability (**Fig. S8**) and thus amino acid positions 106, 115 for A110D and 181, 202 for S189A were randomly mutated (NNK).

Plasmid Construction

To construct the libraries, we first identified naturally occurring, or, when necessary, inserted restriction sites that were in close proximity to the desired S.C. and mutagenesis sites within the S. cerevisiae HIS3 gene. Single-stranded oligonucleotides that coded for the desired amino acid mutations (including all synonymous variants) and were flanked by the selected restriction sites plus 15-25bp homology were designed accordingly.

Genomic DNA was extracted (Multi Sciences, MK010001) from the prototrophic S288C derived FY4 strain. The HIS3 locus (ATG - 250bp, TAG + 300bp) was then amplified using 500ng of the extracted DNA, primers 376, 379, and the Golden Star T6 DNA polymerase (Tsingke Biotech, TSE101). For S189A, the amplicon was PCR purified (Qiagen MinElute, 28004) and inserted between the SacI and KpnI sites (destroying them) within the pGREG506 ADH1pr-MCM-mCherry vectors (a gift from the Chao Tang lab) via Gibson Assembly (Tiangen Easygeno, VI201) to yield plasmid p142. The A110D library required the insertion of an EagI site which was accomplished using two additional primers (primers 342, 343) that contained a 5” overlap and remaining EagI sequence (GGCCG preceded by a C). As with S189A, post PCR both amplicons were verified, purified, and assembled into the vector backbone to yield plasmid p143.

Library Construction

To construct the final libraries, 7.5µg of p142 and p143 were digested for 4 hours at 37°C with 9µL of KpnI, NheI (NEB, R3131S), and BstXI (NEB, R01131S) and EagI (NEB, R3505S) respectively. Each reaction was then loaded on a 0.7% agarose gel, with 0.5x TAE running buffer, gel purified (Magen, D2111), and concentrated (Tianmo Biotech, TD413). Subsequently, the library ssOligos (primers 337, 338) were then assembled into 500ng of the digested vectors to a final concentration of 0.25µM in 15µL Gibson Assembly reactions (2 step Gibson Assembly -3 minutes at 37°C, 60 minutes at 50°C). Finally, the assembly reactions were split into 5 transformations in chemically competent E. Coli (Transgen Biotech, CD101) and spread onto 15 LB+Amp plates, and incubated overnight (~14 hours) at 37°C. Post incubation, the colonies were verified (colony PCR, Sanger sequencing) and scraped off the plates (~55,000 colonies/library), mixed and the library plasmids were extracted (Tiangen Biotech, DP106), quantified, and stored.

Yeast Transformation and Bulk Competition

Competent yeast cells were prepared (Zymo Research, T2001) using yeast strain Y120 (derived from UCC8363; MKT1-RM SAL1-RM CAT5-RM MIP1-BY ura3Δ0 his3::KanMX). Subsequently, 5µg of each plasmid was transformed into ~52.5 x 10^6 cells and incubated at 30°C, 220RPM for 50hrs in 1L synthetic complete dropout medium (SCD) lacking Uracil (-Ura); samples (0.5, 5.5µL) were also spread on SCD-Ura agar plates and incubated at 30°C for 50 hrs.
Post transformation (yielding ~750,000 colonies/library), ~3 x 10^8 cells per library were inoculated into 1L of SCD-Ura and grown at 30°C, 220RPM for 6hrs. Post-pre-growth, ~3 x 10^7 cells per library were inoculated into 1L SCD-His (Histidine) and ~2.1 x 10^7 cells into 700mL SCD-Ura and incubated at 30°C, 220RPM. The competition was carried out for 120 hours with 12 hours between bottlenecks (3 and 2.1 x 10^7 cells per library transferred into fresh media), two 15mL samples (technical replicates) were taken from each culture every 24hrs, centrifuged (3000g, 5 mins) and stored (-20°C), the experiment from transformation to completion was carried out twice (biological replicates).

Plasmid extraction and Illumina sequencing.

Post competition, cell pellets from each time point were thawed and plasmids extracted using the Tiangen Midiprep kit (DP106). Extracted DNA (~10ng) was labeled and amplified using staggered primers (primers 345-356) and Q5 DNA polymerase (NEB, M0492S) in 50µL reactions (98°C for the 30s; 98°C for 10s, 55°C for 30s and 72°C for 30s for 17 cycles; 72°C for 2 min). The amplicons were purified (Qiagen MinElute, 28004) and eight pooled samples (S189A and A110D samples mixed in an equimolar ratio) were submitted to Genewiz for Illumina library construction (A tailed adapter ligation) and sequencing.

Data processing, and per-A.A. genotype fitness calculation.

Raw paired-end FASTQ files were merged using FLASH (Magoc and Salzberg, 2011), demultiplexed according to the unique primer pairs used to amplify each sample and timepoint, and reads for each genotype counted using starcode (Zorita et al., 2015). Reads for each A.A. genotype were the sum of all synonymous NT genotypes and the sum across two technical replicates. Fitness for each A.A genotype was calculated using PyFitSeq (Li et al., 2018), and the average fitness across both biological replicates was used for further analysis.

Competing interests

The authors have declared that no competing interests exist.

Author contributions.

W.H.J., S.S.K. and H.C.Y. supervised the study. W.H.J., S.S.K., H.C.Y. and Yiparemu carried out the experiments, analyzed the results, plotted the figures, and wrote the manuscript. All the authors read and accepted the final version of the manuscript.

Acknowledgments

We'd like to thank Dr.Chu Wang and Dr. Luhua Lai for helpful discussions and comments on the manuscript.

References

Alauddin M, Biswal HS, Gloaguen E, Mons M. 2015. Intra-residue interactions in proteins: interplay between serine or cysteine side chains and backbone conformations, revealed by laser spectroscopy of isolated model peptides. Phys Chem Chem Phys 17: 2169-78. https://doi.org/10.1039/c4cp04449e

Bak P, Flyvbjerg H, Lautrup B. 1992. Coevolution in a rugged fitness landscape. Phys Rev A 46: 6724-6730. https://doi.org/10.1103/physreva.46.6724

Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS. 2006. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. Nature 444: 929-32. https://doi.org/10.1038/nature05385

Bloom JD, Gong LI, Baltimore D. 2010. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. Science 328: 1272-5. https://doi.org/10.1126/science.1187816

Bloom JD, Lu Z, Chen D, Raval A, Venturelli OS, Arnold FH. 2007. Evolution favors protein mutational robustness in sufficiently large populations. BMC Biol 5: 29. https://doi.org/10.1186/1741-7007-5-29

Boeke JD, Trueheart J, Natsoulis G, Fink GR. 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol 154: 164-75. https://doi.org/10.1016/0076-6879(87)54076-9

Casadio R, Vassura M, Tiwari S, Fariselli P, Luigi Martelli P. 2011. Correlating disease-related mutations to their effect on protein stability: a large-scale analysis of the human proteome. Hum Mutat 32: 1161-70. https://doi.org/10.1002/humu.21555

Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan J, Niemann S, Gagneux S. 2011. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. Nat Genet 44: 106-10. https://doi.org/10.1038/ng.1038

de Visser JAGM, Elena SF, Fragata I, Matuszewski S. 2018. The utility of fitness landscapes and big data for predicting evolution. Heredity (Edinb) 121: 401-405. https://doi.org/10.1038/s41437-018-0128-4

Domingo J, Diss G, Lehner B. 2018. Pairwise and higher-order genetic interactions during the evolution of a tRNA. Nature 558: 117-121. https://doi.org/10.1038/s41586-018-0170-7

Fang X, Huang J, Zhang R, Wang F, Zhang Q, Li G, Yan J, Zhang H, Yan Y, Xu L. 2019. Convolution Neural Network-Based Prediction of Protein Thermostability. J Chem Inf Model 59: 4833-4843. https://doi.org/10.1021/acs.jcim.9b00220

Faul F, Erdfelder E, Buchner A, Lang AG. 2009. Statistical power analyses using GPower 3.1: tests for correlation and regression analyses. Behav Res Methods 41: 1149-60. https://doi.org/10.3758/BRM.41.4.1149

Faure G, Koonin EV. 2015. Universal distribution of mutational effects on protein stability, uncoupling of protein robustness from sequence evolution and distinct evolutionary modes of prokaryotic and eukaryotic proteins. Phys Biol 12: 035001. https://doi.org/10.1088/1478-3975/12/3/035001

Fay JC, Wyckoff GJ, Wu CI. 2001. Positive and negative selection on the human genome. Genetics 158: 1227-34. https://doi.org/10.1093/genetics/158.3.1227

Ferrer-Costa C, Orozco M, de la Cruz X. 2007. Characterization of compensated mutations in terms of structural and physico-chemical properties. J Mol Biol 365: 249-56. https://doi.org/10.1016/j.jmb.2006.09.053

Filteau M, Hamel V, Pouliot MC, Gagnon-Arsenault I, Dubé AK, Landry CR. 2015. Evolutionary rescue by compensatory mutations is constrained by genomic and environmental backgrounds. Mol Syst Biol 11: 832. https://doi.org/10.15252/msb.20156444

Handel A, Regoes RR, Antia R. 2006. The role of compensatory mutations in the emergence of drug resistance. PLoS Comput Biol 2: e137. https://doi.org/10.1371/journal.pcbi.0020137

Hanson DK, Tiede DM, Nance SL, Chang CH, Schiffer M. 1993. Site-specific and compensatory mutations imply unexpected pathways for proton delivery to the QB binding site of the

photosynthetic reaction center. Proc Natl Acad Sci U S A 90: 8929-33. https://doi.org/10.1073/pnas.90.19.8929

Huang LT, Saraboji K, Ho SY, Hwang SF, Ponnuswamy MN, Gromiha MM. 2007. Prediction of protein mutant stability using classification and regression tool. Biophys Chem 125: 462-70. https://doi.org/10.1016/j.bpc.2006.10.009

Ivankov DN, Finkelstein AV, Kondrashov FA. 2014. A structural perspective of compensatory evolution. Curr Opin Struct Biol 26: 104-12. https://doi.org/10.1016/j.sbi.2014.05.004

Jakubowska A, Korona R. 2009. Lack of evolutionary conservation at positions important for thermal stability in the yeast ODCase protein. Mol Biol Evol 26: 1431-4. https://doi.org/10.1093/molbev/msp066

Jaspard E, Hunault G. 2014. Comparison of amino acids physico-chemical properties and usage of late embryogenesis abundant proteins, hydrophilins and WHy domain. PLoS One 9: e109570. https://doi.org/10.1371/journal.pone.0109570

Jelesarov I, Karshikoff A. 2009. Defining the role of salt bridges in protein stability. Methods Mol Biol 490: 227-60. https://doi.org/10.1007/978-1-59745-367-7_10

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22: 2577-637. https://doi.org/10.1002/bip.360221211

Kawashima S, Pokarowski P, Pokarowska M, Kolinski A, Katayama T, Kanehisa M. 2008. AAindex: amino acid index database, progress report 2008. Nucleic Acids Res 36: D202-5. https://doi.org/10.1093/nar/gkm998

Kim HW, Shen TJ, Sun DP, Ho NT, Madrid M, Tam MF, Zou M, Cottam PF, Ho C. 1994. Restoring allosterism with compensatory mutations in hemoglobin. Proc Natl Acad Sci U S A 91: 11547-51. https://doi.org/10.1073/pnas.91.24.11547

Kimura M. 1985. The role of compensatory neutral mutations in molecular evolution. Journal of Genetics 64: 7-19.

Kirby DA, Muse SV, Stephan W. 1995. Maintenance of pre-mRNA secondary structure by epistatic selection. Proc Natl Acad Sci U S A 92: 9047-51. https://doi.org/10.1073/pnas.92.20.9047

Kondrashov AS, Sunyaev S, Kondrashov FA. 2002. Dobzhansky-Muller incompatibilities in protein evolution. Proc Natl Acad Sci U S A 99: 14878-83. https://doi.org/10.1073/pnas.232565499

Kondrashov DA, Kondrashov FA. 2015. Topological features of rugged fitness landscapes in sequence space. Trends Genet 31: 24-33. https://doi.org/10.1016/j.tig.2014.09.009

Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, Desai MM. 2013. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500: 571-4. https://doi.org/10.1038/nature12344

Lazar T, Guharoy M, Schad E, Tompa P. 2018. Unique Physicochemical Patterns of Residues in Protein-Protein Interfaces. J Chem Inf Model 58: 2164-2173. https://doi.org/10.1021/acs.jcim.8b00270

Li F, Salit ML, Levy SF. 2018. Unbiased fitness estimation of pooled barcode or amplicon sequencing studies. Cell Syst 7: 521-525.e4. https://doi.org/10.1016/j.cels.2018.09.004

Lin CP, Huang SW, Lai YL, Yen SC, Shih CH, Lu CH, Huang CC, Hwang JK. 2008. Deriving protein dynamical properties from weighted protein contact number. Proteins 72: 929-35. https://doi.org/10.1002/prot.21983

Lind PA, Arvidsson L, Berg OG, Andersson DI. 2017. Variation in Mutational Robustness between Different Proteins and the Predictability of Fitness Effects. Mol Biol Evol 34: 408-418. https://doi.org/10.1093/molbev/msw239

Liu J, Kang X. 2012. Grading amino acid properties increased accuracies of single point mutation on protein stability prediction. BMC Bioinformatics 13: 44. https://doi.org/10.1186/1471-2105-13-44

Liu Q, Zuo T, Xu P, Jiang Q, Wu J, Gan M, Yang C, Prakash R, Zhu G, Takiff HE, et al. 2018. Have compensatory mutations facilitated the current epidemic of multidrug-resistant tuberculosis? Emerg Microbes Infect 7: 98. https://doi.org/10.1038/s41426-018-0101-6

Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27: 2957-63. https://doi.org/10.1093/bioinformatics/btr507

Maisnier-Patin S, Andersson DI. 2004. Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. Res Microbiol 155: 360-9. https://doi.org/10.1016/j.resmic.2004.01.019

Mateu MG, Fersht AR. 1999. Mutually compensatory mutations during evolution of the tetramerization domain of tumor suppressor p53 lead to impaired hetero-oligomerization. Proc Natl Acad Sci U S A 96: 3595-9. https://doi.org/10.1073/pnas.96.7.3595

Meftahi N, Namouchi A, Mhenni B, Brandis G, Hughes D, Mardassi H. 2016. Evidence for the critical role of a secondary site rpoB mutation in the compensatory evolution and successful transmission of an MDR tuberculosis outbreak strain. J Antimicrob Chemother 71: 324-32. https://doi.org/10.1093/jac/dkv345

Morton NE, Crow JF, Muller HJ. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. Proc Natl Acad Sci U S A 42: 855-63. https://doi.org/10.1073/pnas.42.11.855

Narayanan C, Dias CL. 2013. Hydrophobic interactions and hydrogen bonds in β-sheet formation. J Chem Phys 139: 115103. https://doi.org/10.1063/1.4821596

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205-17. https://doi.org/10.1006/jmbi.2000.4042

Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. 2007. Crystal structure of an ancient protein: evolution by conformational epistasis. Science 317: 1544-8. https://doi.org/10.1126/science.1142819

Palmer AC, Toprak E, Baym M, Kim S, Veres A, Bershtein S, Kishony R. 2015. Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. Nat Commun 6: 7385. https://doi.org/10.1038/ncomms8385

Pokusaeva VO, Usmanova DR, Putintseva EV, Espinar L, Sarkisyan KS, Mishin AS, Bogatyreva NS, Ivankov DN, Akopyan AV, Avvakumov SY, et al. 2019. An experimental assay of the interactions of amino acids from orthologous sequences shaping a complex fitness landscape. PLoS Genet 15: e1008079. https://doi.org/10.1371/journal.pgen.1008079

Poon A, Chao L. 2005. The rate of compensatory mutation in the DNA bacteriophage phiX174. Genetics 170: 989-99. https://doi.org/10.1534/genetics.104.039438

Poon AF, Chao L. 2006. Functional origins of fitness effect-sizes of compensatory mutations in the DNA bacteriophage phiX174. Evolution 60: 2032-43.

Rawson S, Bisson C, Hurdiss DL, Fazal A, McPhillie MJ, Sedelnikova SE, Baker PJ, Rice DW, Muench SP. 2018. Elucidating the structural basis for differing enzyme inhibitor potency by cryo-EM. Proc Natl Acad Sci U S A 115: 1795-1800. https://doi.org/10.1073/pnas.1708839115

Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Ivankov DN, Bozhanova NG, Baranov MS, Soylemez O, et al. 2016. Local fitness landscape of the green fluorescent protein. Nature 533: 397-401. https://doi.org/10.1038/nature17995

Sengupta D, Kundu S. 2012. Role of Physico-chemical Properties of Amino Acids in Protein's Structural Organization: A Network Perspective. In Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28792-3_11

Stein A, Fowler DM, Hartmann-Petersen R, Lindorff-Larsen K. 2019. Biophysical and Mechanistic Models for Disease-Causing Protein Variants. Trends Biochem Sci 44: 575-588. https://doi.org/10.1016/j.tibs.2019.01.003

Stephan W, Kirby DA. 1993. RNA folding in Drosophila shows a distance effect for compensatory fitness interactions. Genetics 135: 97-103. https://doi.org/10.1093/genetics/135.1.97

Storz JF. 2018. Compensatory mutations and epistasis for protein function. Curr Opin Struct Biol 50: 18-25. https://doi.org/10.1016/j.sbi.2017.10.009

Szendro IG, Franke J, de Visser JA, Krug J. 2013. Predictability of evolution depends nonmonotonically on population size. Proc Natl Acad Sci U S A 110: 571-6. https://doi.org/10.1073/pnas.1213613110

Teng S, Srivastava AK, Wang L. 2010. Sequence feature-based prediction of protein stability changes upon amino acid substitutions. BMC Genomics 11 Suppl 2: S5. https://doi.org/10.1186/1471-2164-11-S2-S5

Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. 2013. Maximum allowed solvent accessibilities of residues in proteins. PLoS One 8: e80635. https://doi.org/10.1371/journal.pone.0080635

Tokuriki N, Tawfik DS. 2009. Stability effects of mutations and protein evolvability. Curr Opin Struct Biol 19: 596-604. https://doi.org/10.1016/j.sbi.2009.08.003

Tomala K, Zrebiec P, Hartl DL. 2019. Limits to Compensatory Mutations: Insights from Temperature-Sensitive Alleles. Mol Biol Evol 36: 1874-1883. https://doi.org/10.1093/molbev/msz110

Touw WG, Baakman C, Black J, te Beek TA, Krieger E, Joosten RP, Vriend G. 2015. A series of PDB-related databanks for everyday needs. Nucleic Acids Res 43: D364-8. https://doi.org/10.1093/nar/gku1028

Van Cleve J, Weissman DB. 2015. Measuring ruggedness in fitness landscapes. Proc Natl Acad Sci U S A 112: 7345-6. https://doi.org/10.1073/pnas.1507916112

Wu NC, Dai L, Olson CA, Lloyd-Smith JO, Sun R. 2016. Adaptation in protein fitness landscapes is facilitated by indirect paths. Elife 5: e16965. https://doi.org/10.7554/eLife.16965

Zeldovich KB, Chen P, Shakhnovich EI. 2007. Protein stability imposes limits on organism complexity and speed of molecular evolution. Proc Natl Acad Sci U S A 104: 16152-7. https://doi.org/10.1073/pnas.0705366104

Zheng J, Payne JL, Wagner A. 2019. Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. Science 365: 347-353. https://doi.org/10.1126/science.aax1837

Zorita E, Cuscó P, Filion GJ. 2015. Starcode: sequence clustering based on all-pairs search. Bioinformatics 31: 1913-19. https://doi.org/10.1093/bioinformatics/btv053

Figures and Tables

Figure 1. Widespread compensatory substitutions in an evolutionarily-relevant deep mutational scanning library. (A) Compensatory substitutions are additional substitutions that rescue loss-of-function genotypes. **(B)** An example of compensatory substitutions in which a genotype with the fitness of zero (blue) is four substitutions away from wild-type and for which we measured the fitness of 23 genotypes with additional substitutions. The green lines connect to rescuing genotypes, and the red lines connect to non-rescuing genotypes. **(C)** Quantification of the example from (B) shows how to calculate the rescuability. The blue dot shows the unfit genotype for which we calculate rescuability and the black dots are genotypes which fail to rescue. Error bars are the standard deviation in the estimation of fitness. **(D)** Total counts of unfit genotypes grouped by their rescuability scores. **(E)** The distribution of rescuability of unfit genotypes (purple for all missense genotypes and blue for missense genotypes with N >= 10). The green bars show the low rescuability of nonsense genotypes as a control.

Figure 2 Rescuability is correlated with structural features and evolutionary conservation. (A) The rescuability of each position in the monomer of His3p was calculated as the average rescuability of all genotypes with that position substituted. Error bars are standard deviation. Colors represent the type of secondary structure. Nonsense

genotypes (grey) are not expected to be rescuable and are shown as a control. Positions which are located at interface regions are shown in black. **(B)** The mean rescuability is significantly higher at loop positions (Wilcoxon rank-sum test with one tail) than α-helix and β-strand positions. **(C)** The rescuability of each position is correlated with structural features and conservation, which are all correlated with each other. RSA: relative solvent accessibility, WCN: weighted contact number. **(D)** Genotypes in which more than 50% of substitutions are in loop regions are more rescuable (t-test with one tail), error bars are standard deviation from bootstrapping).

Figure 3. Rescuable genotypes are more structurally stable than non-rescuable genotypes, and rescuing genotypes can compensate for the stability loss of rescuable genotypes. (A) Rescuable genotypes have smaller △△G of protein structure than non-rescuable genotypes. These genotypes are one subset with a different number of substitutions (Wilcoxon rank-sum test with one tail). **(B)** Most rescuing genotypes have smaller △△G of protein structure than rescuable genotypes. **(C)** For genotypes with more than 50% substitutions in interface regions, rescuable genotypes have smaller △△G of protein structure than non-rescuable genotypes (Wilcoxon rank-sum test with one tail). **(D)** For genotypes with more than 50% substitutions in interface regions, most rescuing genotypes have smaller △△G of protein structure than rescuable genotypes.

Figure 4. Amino acid physicochemical features are more predictive of fitness in ordered regions. (A) Two examples of physicochemical features CAJ730102(*normalized frequency of β-sheet*) and KARS160109 *(Maximum eigenvalue of the Laplacian matrix of the graph representing the structure of the amino acid)*, in which changes in the value (x-axis: the absolute difference of the sum of the per-amino acid value across all positions in each genotype and the sum in the wild type genotype) are predictive of fitness in specific regions but not in others. **(B)** The protein structure of the substituted positions shown in (A). **(C)** Heatmap of the Spearman correlation between fitness and each amino acid physicochemical feature value (distance to WT) of genotypes with the same set of substituted positions. Columns represent physicochemical features, and rows represent unique sets of mutated positions. The heatmap is clustered by the columns and then clustered rows are sorted by the percentage of substitutions that are in the disordered loop

regions. **(D)** Physicochemical features are more predictive of fitness in ordered regions. Each dot shows the percentage of regions (rows in **C**) in which a single physicochemical feature significantly correlates with fitness (Spearman correlation <= -0.5, p-value<0.05). The threshold of 40% reflects a visual separation observed near the value in (C)

Figure 5. Changes in physicochemical features that predict fitness and rescue. (A) A schematic example to determine if one physicochemical feature can predict rescue in a specific region. The feature can predict rescue when the additional substitutions indicated by the green arrow reduce the feature value's distance from the wild type and enhance fitness. For the further analysis, we determined N1 for the pair of genotypes of which the feature can predict rescue by additional substitutions and N2 for the pair of genotypes of which the feature can't predict rescue. **(B)** Changes in physicochemical features that predict fitness can also predict rescue. Each dot represents one cell from the heatmap in Figure 4C, corresponding to a specific physicochemical feature and its correlation with fitness (Spearman correlation). The x-axis, N1/(N1+N2) *100 (x-axis), quantifies the ability of this physicochemical feature to predict rescue.

Figure 6. Super compensatory substitutions increase fitness in diverse genetic backgrounds. (A) A schematic figure quantifies the compensatory ability of A110D in the 115E background. The compensatory ability of A110D in the 115E background is represented as an arrow that from 115E genotype and ends at A110D. **(B-C)** A few substitutions from wild-type amino acids can increase fitness in diverse genetic backgrounds. (B) shows the whole pairs of one substitution and each amino acid state, like A110D and 115E in (A). The shade of the arrow represents the ratio of genotypes of which fitness increases at least 0.5 with the substitution of A110D. **(C)** A few substitutions have a high compensatory ability. The y-axis represents the compensatory ability of each substitution from the wild-type amino acid, which is quantified as the sum of weights of arrows in (B) which ends at this substitution. Only segment 1 is shown. **(D)** Only a few substitutions have a high compensatory ability in the whole dataset. **(E)** A neural network model predicts fitness by calculating a hidden variance (fitness potential) to map to the fitness of genotypes. And each amino acid state has one weight that we stated as a fitness impact score. **(F)** Substitutions increase fitness by increasing fitness potential. A110D (red)

can increase fitness potential and K115E(green) decreases fitness potential. **(G)** Substitutions that have higher compensatory ability also tend to have higher fitness impact scores than WT.

Figure 7. Super compensators reduce the fitness effects of both deleterious and beneficial substitutions. (A-B) Two examples show that 110D can reduce the fitness effect of K115T (deleterious substitution) and V114I (beneficial substitution). **(C-D, G-H, K)** The super compensator-110D can reduce the fitness effects of substitutions compared with 110A (wild-type amino acid). In (C), the x-axis represents the fitness effects of substitutions with 110D. The y-axis represents the fitness effects of substitutions with 110A. In (D), the x-axis represents the absolute fitness effects of substitutions with 110D and the y-axis represents the absolute fitness effects of substitutions in 110A. (G) shows the distribution of the differences of each one pair of fitness effects of the same beneficial substitution but one with 110D and another with 110A, separately. The positive skewness of this distribution is also shown. (H) shows the distribution of the differences of each one pair of fitness effects of the same deleterious substitution but one with 110D and another with 110A, separately. The positive skewness of this distribution is also shown. (K) shows the distribution of the merged data from (G) and (H). The positive skewness of this distribution is shown. **(E-F, I-J, L)** A non-super compensator-110Q has no effects on the fitness effects of substitutions compared with 110A (wild-type amino acid). (I-J, L) used the same method as (G-H, K) to shows each one pair of fitness effects of the same substitution but one with 110Q and another with 110A, separately.

Figure 8. **Super compensators shift genotypes into a high fitness plateau where the fitness effects of substitutions become less sensitive. (A-B)** Non-WT amino acid states (like 110D) which have positive fitness impact scores than WT amino acid states can move genotypes into a high fitness plateau where fitness effects are insensitive. Non-WT amino acid states which have negative fitness impact scores than WT amino acid states (like 115E) can shift genotypes away from the high fitness plateau and make fitness effects sensitive (F.E.: fitness effect.) (A) shows an example that super compensator 93V can reduce fitness effects of substitutions and make the local fitness landscape flatter. **(C)** The distribution of the absolute differences of each one pair of fitness effects of the same

substitution but one with 110A and another with 110D shows positive skewness. **(D)**The distribution of the absolute differences of each one pair of fitness effects of the same substitution but one with 115K and another with 115E has negative skewness. **(E)** Non-WT amino acid states with higher compensatory ability exhibit stronger buffering ability, as evidenced by more positive skewness and thus reduce the fitness effects of substitutions (Spearman R= 0.67, p-value = 0). **(F)** Non-WT amino acid states with higher positive fitness impact scores relative to WT amino acid states also exhibit stronger buffering ability, as shown by higher positive skewness and thus reduce fitness effects of substitutions (Spearman R= 0.78, p-value = 0).

Figure 9. Super compensators co-exist with other deleterious amino acid states in the extant species. (**A**) Super compensators co-exist with other amino acid states in the extant 355 species. Arrows represent co-occurrence relationships between amino acid states, with the shade of the arrow indicating the fraction of species in which the pair co-exists. Pairs of amino acid states that co-exist in at least 10 species are labeled in white. (**B**) Among pairs of non-WT amino acid states in the extant species, if one amino acid state tends to exhibit a low fitness impact score relative to WT amino acid states, another amino acid state tends to has a high fitness impact score.

Figure 10. Super compensators increase protein local structurally stability. Non-WT amino acid states with higher fitness impact scores relative to WT amino acid states are associated with more stable protein structures, as reflected by lower $\triangle\triangle G$ energy.

Figure 11. Super compensators of 189A and 110D increase the fitness of genotypes with random mutations comparing with the wild type amino acid states in the same positions. (A) The amino acids frequency of the constructed library shows that the super compensator 189A increases the fitness of genotypes with random mutations compared to the wild-type amino acid state 189S. Specifically, 18 genotypes are fit with 189A but unfit with 189S, while only 3 genotypes are fit with 189S but unfit with 189A. **(B-C)** The genotypes with stop codons have low fitness in the absence of histidine (left panel, orange) comparing to the presence of histidine (right panel, orange), confirming the reliability of the built library. However, only a small fraction of genotypes in absence of stop codons

are unfit in absence of histidine. **(D)** The average fitness difference between genotypes with 189A and with 189S is higher than zero, and the fitness difference is more pronounced in deleterious genotypes (orange, Wilcoxon-sum test with one tail). (D)The average fitness difference between genotypes with 110D and with 110A is higher than zero, and the fitness difference is more evident in deleterious genotypes (orange, Wilcoxon-sum test with one tail).

Figure S1. The library is constructed based on extant amino acids states and covers almost the entire length of His3p. (A) The red heatmap shows the conservation score among 21 extant yeasts at each position of the monomer of **His3p**, while the blue heatmaps show the conservation score and the frequency of each amino acid in each position in the library. **(B)** Most genotypes are composed of extant amino acids. **(C)** The fitness distribution of all genotypes in the library. **(D)** Percentage of unfit genotypes grouped by number of substitutions contained. (**E**) Unfit genotypes cover almost the entire length of **His3p**. **(F)** The rescuability of each unfit genotype is little correlated with the number of substitutions contained (Pearson R= 0.18, p:3.2e-239).

Figure S2. Different rescuability calculation methods and filtering thresholds produce similar results. (A-F) Different filtering thresholds used to calculate the average rescuability for each position of the monomer of **His3p** give similar results as Figure 2. **(G)** Different rescuability calculation methods produce similar rescuability for unfit genotypes ($R^2 = 0.84$).

FigureS3.The correlation between rescuability and secondary structure features of protein is robust. (A-B) The average rescuability in loop regions is significantly higher than in β-strand regions after removing positions with average rescuability greater than 0.75 (D) or greater than 0.33 (E). **(C)** The average rescuability is higher in loop regions than in α-helix and β-strand in the protein GB1. (Wilcoxon-sum test with one tail)

FigureS4. The RSA scores are higher in the surface and the C-terminal loop regions relative to other regions, but the WCN as well as conservation scores are lower in these two loops relative to other regions. (A) The position and the conformation of the surface loop and the C-terminal loop in the monomer of **His3p**. The surface loop is in red

and C-terminal loop is in pink. **(B)** The distribution of conservation scores in the monomer of **His3p**. **(C)** The distribution of WCN scores in the monomer of **His3p**. **(D)** The distribution of RSA scores in the monomer of **His3p**. WCN (weighted contact number), RSA (Relative solvent accessibility).

FigureS5. Rescuable genotypes are more structurally stable than non-rescuable genotypes. (A) Most rescuable genotypes are more structurally stable than non-rescuable genotypes with a different number of substitutions. **(B)** For genotypes with more than 50% substitutions in interface regions, the rescuable genotypes are more structurally stable than non-rescuable genotypes with a different number of substitutions.

Figure S6. Changes in physicochemical features can't predict fitness using data with fitness randomly shuffled.

Figure S7. The highly correlated physicochemical features cause the positive correlation between the fitness and the absolute difference of the AAindex values to WT. (A) The PUNT03101 is highly predictive of fitness. **(B)** The WARP780101 is highly predictive of fitness but WT. **(C)** The PUNT03101 and WARP780101 are highly correlated.

Figure S8. There exist several substitutions that rescue diverse genotypes. (A) The compensatory results of every pair of one substitution from WT amino acid states and another non-WT amino acid state in all 12 segments except for segment9 and segment1. **(B)**The compensatory ability of each substitution from WT amino acid states in all 12 segments except for segment9 and segment1.

Figure S9. Fitness potential is a hidden variance to predict fitness in the 12 segments except for segment10.

Figure S10. In the random library of avGFP, there are some super compensator-like amino acid states. (A) The fitness impact scores of non-WT amino acid states from the neural network model are highly correlated with the brightness of variants with one substitution from the wild type amino acid. **(B)** The brightness of variants with 161A (161A has the highest fitness impact score) is higher than variants with 161V (the wild type amino

acid). Each dot represents one pair of genotypes, and the only difference between these two genotypes is that one genotype has 161A and another genotype has 161V. **(C)** The brightness of variants with 218P (218P has the lowest fitness impact score) is lower than variants with 218L (the wild type amino acid). Each dot represents one pair of genotypes, and the only difference between these two genotypes is that one genotype has 218P and another genotype has 218L.

Figure S11. The whole co-exists amino acid states pairs in the extant species in the whole segments except segment 1(Figure 9A), segment 9, and segment10. The non-extant amino acid states are colored in light.

Figure S12. The distance between super compensatory and non-super compensator amino acid states in all 12 segments except for segment 9 and segment 10.

Table S1. the structural effects of top7 super compensators in and the structural effects of 5 non-super compensators are also shown in green.

Table S1: Summary of mutations and their corresponding effects on protein structure

Mutation	Reason	Super compensatory substitutions
S189A	Remarkably stable helical structure	Yes
E169R	Hydrogen bonds that may form between two side chains adjacent to the helix with E165 to stabilize the helix	Yes
N137G	The conformation interval of R138 is enlarged, and the salt bridge between R138 and phosphate group in substrate is weakened	Yes
A110D	Form a hydrogen bond with R112 to stabilize loop conformation	Yes
E144N	To weaken its rejection with D178 and stabilize the beta sheet structure, N may also form hydrogen bonds with D178	Yes
S73A	Remarkably stable helical structure, favorable hydrophobic accumulation	Yes
T93V	It is advantageous for hydrophobic accumulation in helical beam	Yes
I19D	Destroy the beta sheet and helical stack	No
I19N	Destroy the beta sheet and helical stack	No
I200P	Destroy helix, destroy hydrophobic accumulation	No
H67P	Destroy the helix, destroy the hydrogen bond and proton transport	No
F163I	Destroy the helix, destroy the aromatic ring accumulation	No

Figure 1

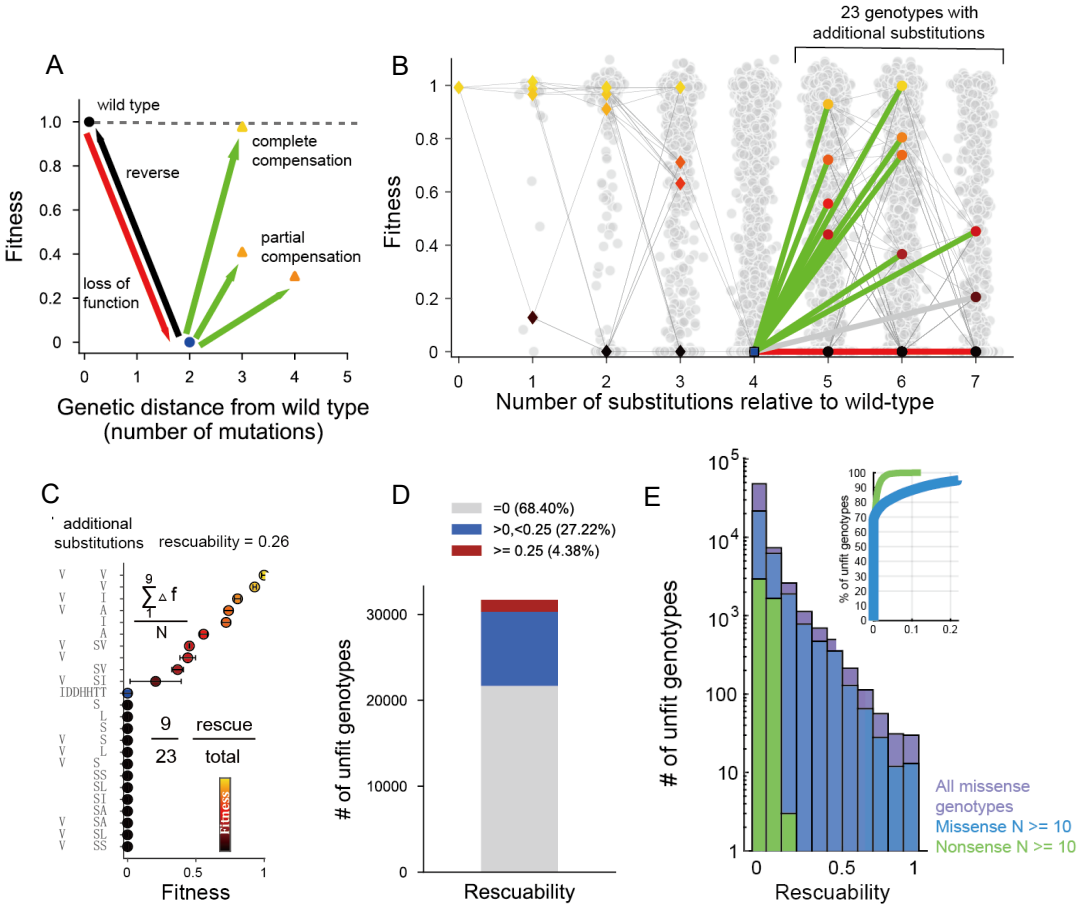

A
wild type
reverse
complete compensation
partial compensation
loss of function
Fitness
Genetic distance from wild type (number of mutations)
B
23 genotypes with additional substitutions
Fitness
Number of substitutions relative to wild-type
C
additional substitutions
rescuability = 0.26
9/23
rescue/total
Fitness
D
=0 (68.40%)
>0,<0.25 (27.22%)
>= 0.25 (4.38%)
# of unfit genotypes
Rescuability
E
# of unfit genotypes
% of unfit genotypes
All missense genotypes
Missense N >= 10
Nonsense N >= 10
Rescuability

Figure 2

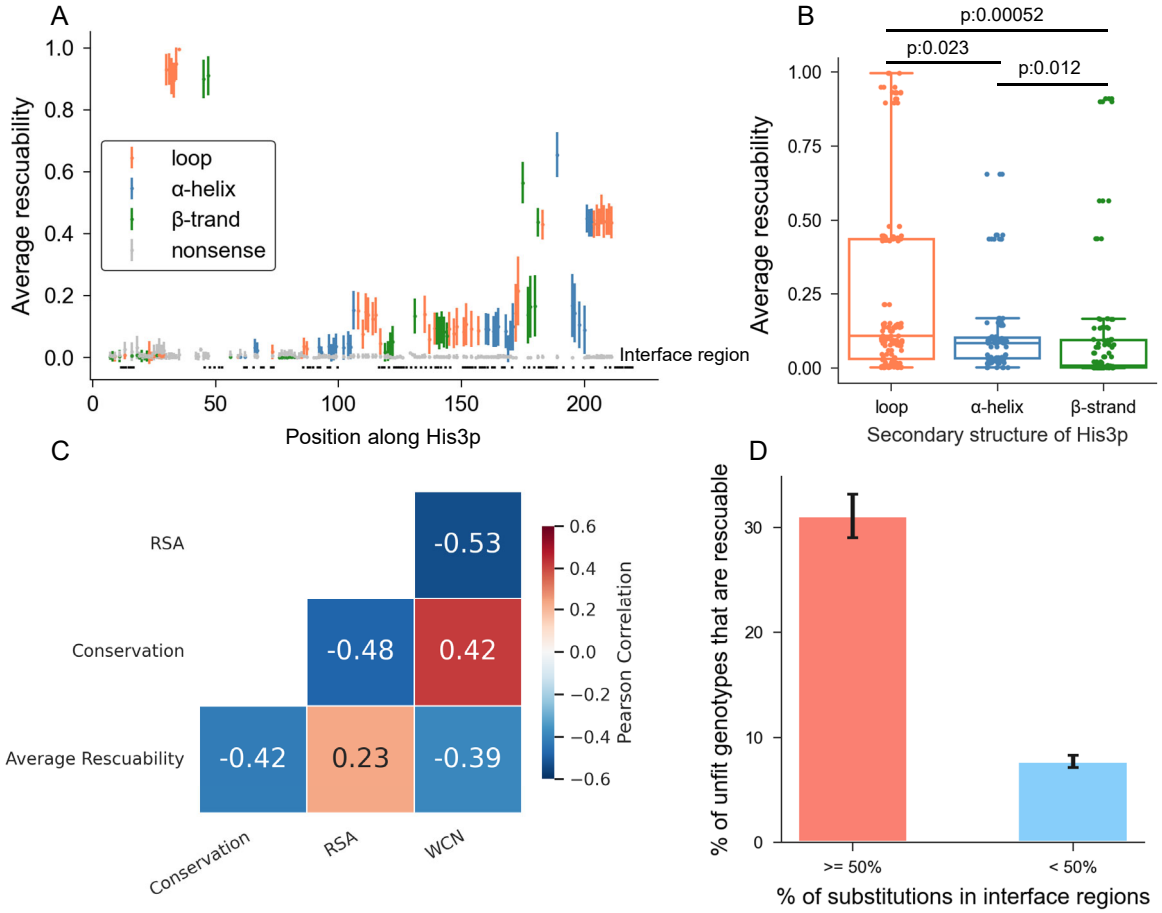

A
Average rescuability
1.0
0.8
0.6
0.4
0.2
0.0
loop
α-helix
β-trand
nonsense
Interface region
0
50
100
150
200
Position along His3p
B
p:0.00052
p:0.023
p:0.012
Average rescuability
1.00
0.75
0.50
0.25
0.00
loop
α-helix
β-strand
Secondary structure of His3p
C
RSA
Conservation
Average Rescuability
-0.53
-0.48
0.42
-0.42
0.23
-0.39
Conservation
RSA
WCN
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Pearson Correlation
D
% of unfit genotypes that are rescuable
30
20
10
0
>= 50%
< 50%
% of substitutions in interface regions

Figure 3

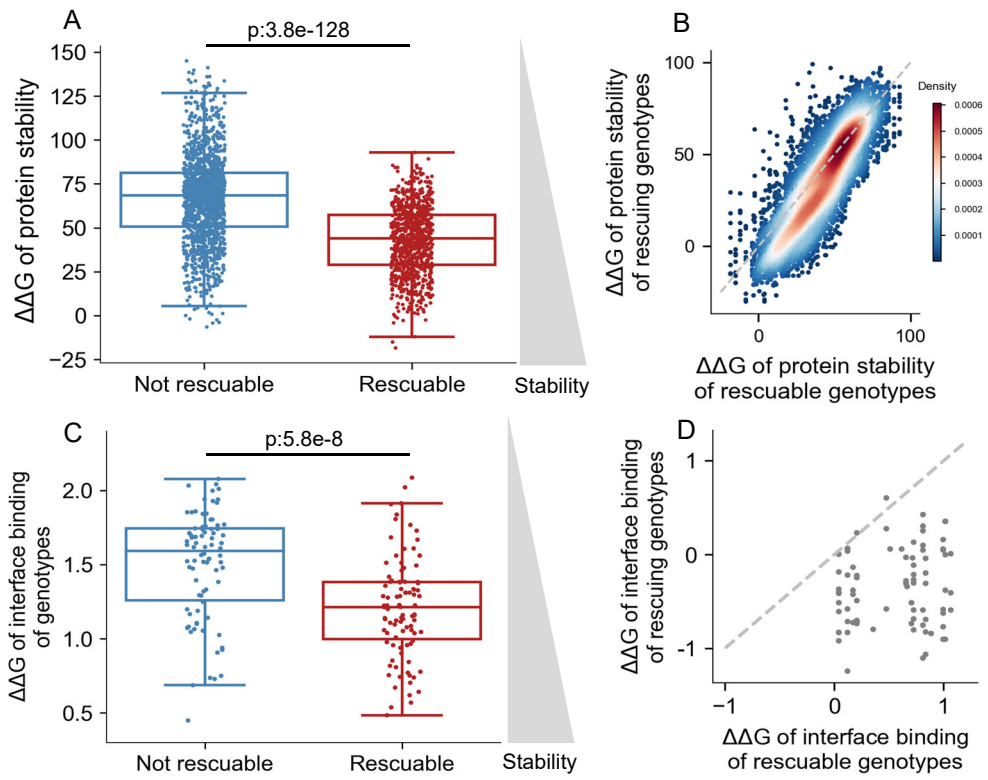

A
p:3.8e-128
ΔΔG of protein stability
150
125
100
75
50
25
0
−25
Not rescuable
Rescuable
Stability
B
ΔΔG of protein stability of rescuing genotypes
100
50
0
0
100
ΔΔG of protein stability of rescuable genotypes
Density
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
C
p:5.8e-8
ΔΔG of interface binding of genotypes
2.0
1.5
1.0
0.5
Not rescuable
Rescuable
Stability
D
ΔΔG of interface binding of rescuing genotypes
1
0
-1
−1
0
1
ΔΔG of interface binding of rescuable genotypes

Figure 4

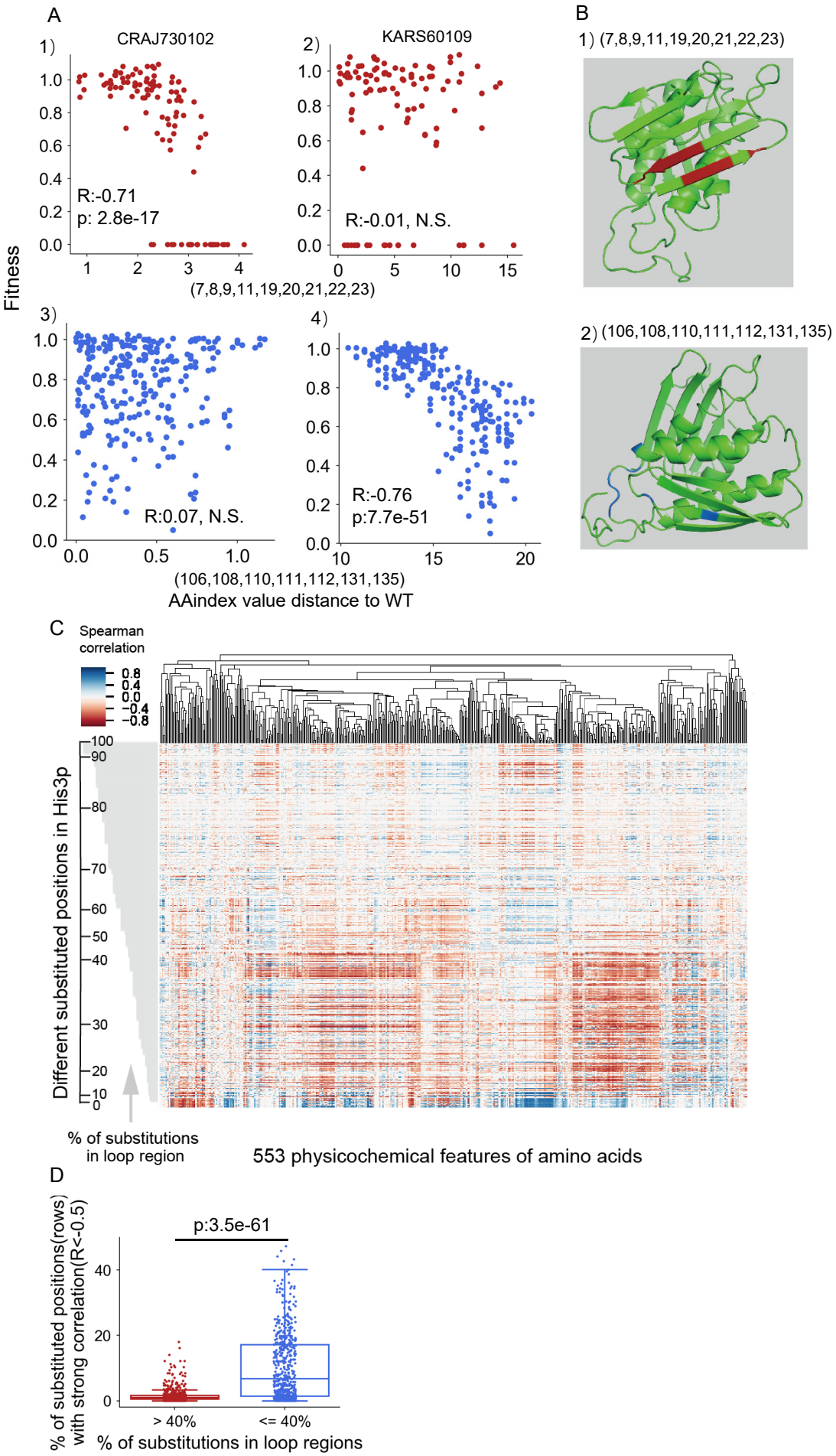

A
1)
CRAJ730102
R:-0.71
p: 2.8e-17
2)
KARS60109
R:-0.01, N.S.
(7,8,9,11,19,20,21,22,23)
Fitness
3)
R:0.07, N.S.
4)
R:-0.76
p:7.7e-51
(106,108,110,111,112,131,135)
AAindex value distance to WT
B
1) (7,8,9,11,19,20,21,22,23)
2) (106,108,110,111,112,131,135)
C
Spearman correlation
Different substituted positions in His3p
% of substitutions in loop region
553 physicochemical features of amino acids
D
p:3.5e-61
% of substituted positions(rows) with strong correlation(R<-0.5)
> 40%
<= 40%
% of substitutions in loop regions

Figure 5

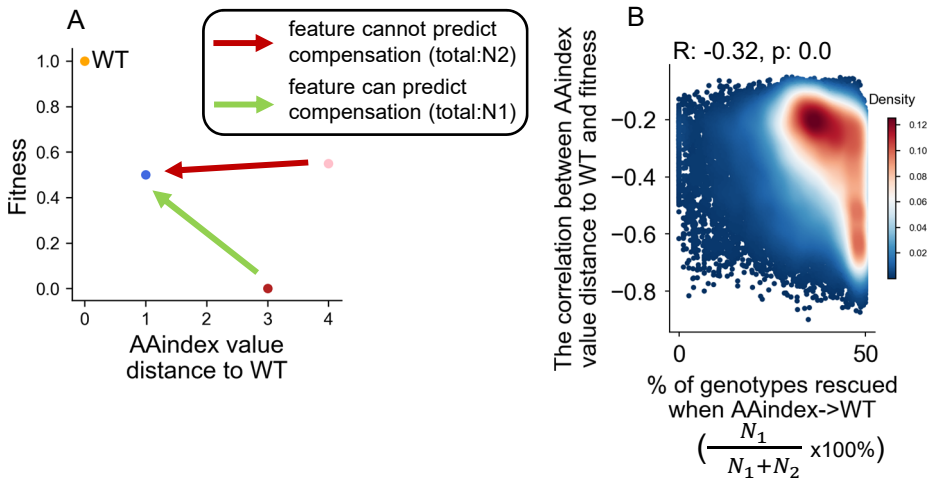

A
WT
feature cannot predict compensation (total:N2)
feature can predict compensation (total:N1)
Fitness
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
AAindex value distance to WT
B
R: -0.32, p: 0.0
Density
0.12
0.10
0.08
0.06
0.04
0.02
−0.2
−0.4
−0.6
−0.8
0
50
The correlation between AAindex value distance to WT and fitness
% of genotypes rescued when AAindex->WT
$\left(\frac{N_1}{N_1+N_2} \text{x}100\%\right)$

Figure 6

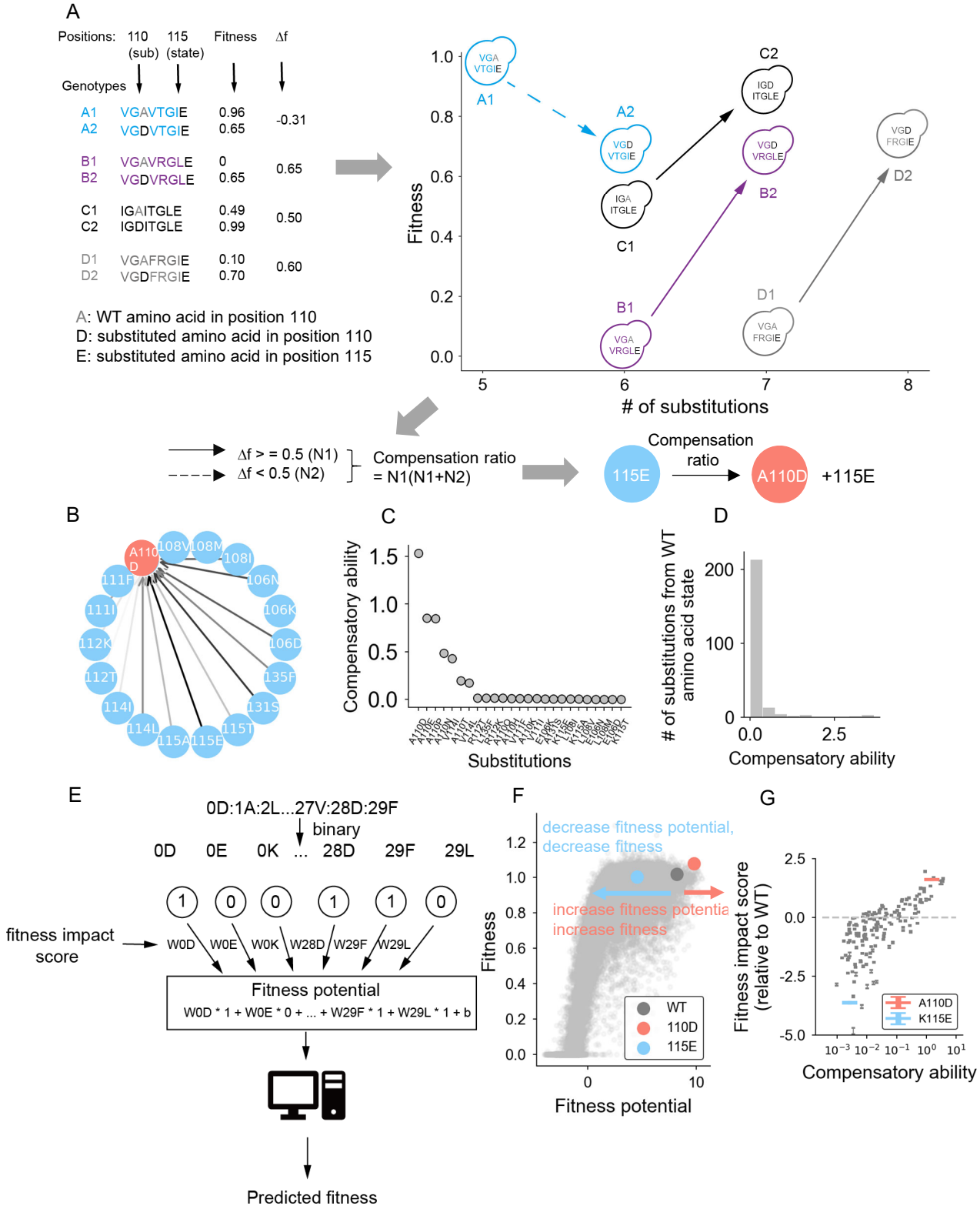

A
Positions: 110 (sub) 115 (state) Fitness Δf
Genotypes
A1 VGAVTGIE 0.96 -0.31
A2 VGDVTGIE 0.65
B1 VGAVRGLE 0 0.65
B2 VGDVRGLE 0.65
C1 IGAITGLE 0.49 0.50
C2 IGDITGLE 0.99
D1 VGAFRGIE 0.10 0.60
D2 VGDFRGIE 0.70
A: WT amino acid in position 110
D: substituted amino acid in position 110
E: substituted amino acid in position 115
Fitness
# of substitutions
Δf > = 0.5 (N1)
Δf < 0.5 (N2)
Compensation ratio = N1(N1+N2)
Compensation ratio
115E
A110D
+115E
B
C
Compensatory ability
Substitutions
D
# of substitutions from WT amino acid state
Compensatory ability
E
0D:1A:2L...27V:28D:29F
binary
0D 0E 0K ... 28D 29F 29L
fitness impact score
W0D W0E W0K W28D W29F W29L
Fitness potential
W0D * 1 + W0E * 0 + ... + W29F * 1 + W29L * 1 + b
Predicted fitness
F
decrease fitness potential, decrease fitness
increase fitness potentia
increase fitness
WT
110D
115E
Fitness
Fitness potential
G
Fitness impact score (relative to WT)
A110D
K115E
Compensatory ability

Figure 7

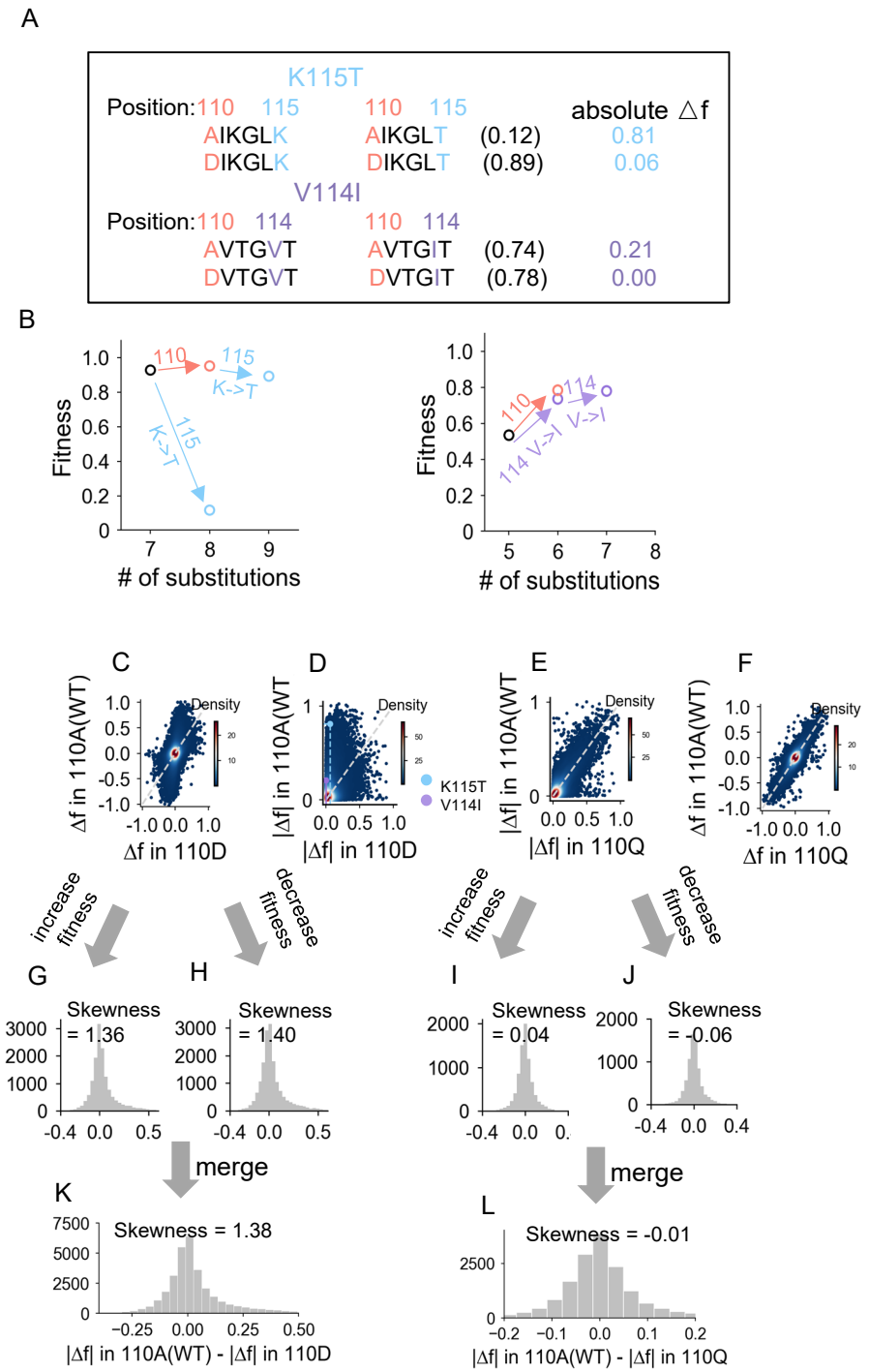

A
K115T
Position: 110 115
110 115
absolute △f
AIKGLK
AIKGLT
(0.12)
0.81
DIKGLK
DIKGLT
(0.89)
0.06
V114I
Position: 110 114
110 114
AVTGVT
AVTGIT
(0.74)
0.21
DVTGVT
DVTGIT
(0.78)
0.00
B
Fitness
# of substitutions
K->T
V->I
C
D
E
F
Density
Δf in 110A(WT)
Δf in 110D
|Δf| in 110A(WT
|Δf| in 110D
|Δf| in 110Q
Δf in 110Q
K115T
V114I
increase fitness
decrease fitness
G
H
I
J
Skewness = 1.36
Skewness = 1.40
Skewness = 0.04
Skewness = -0.06
merge
K
L
Skewness = 1.38
Skewness = -0.01
|Δf| in 110A(WT) - |Δf| in 110D
|Δf| in 110A(WT) - |Δf| in 110Q

Figure 8

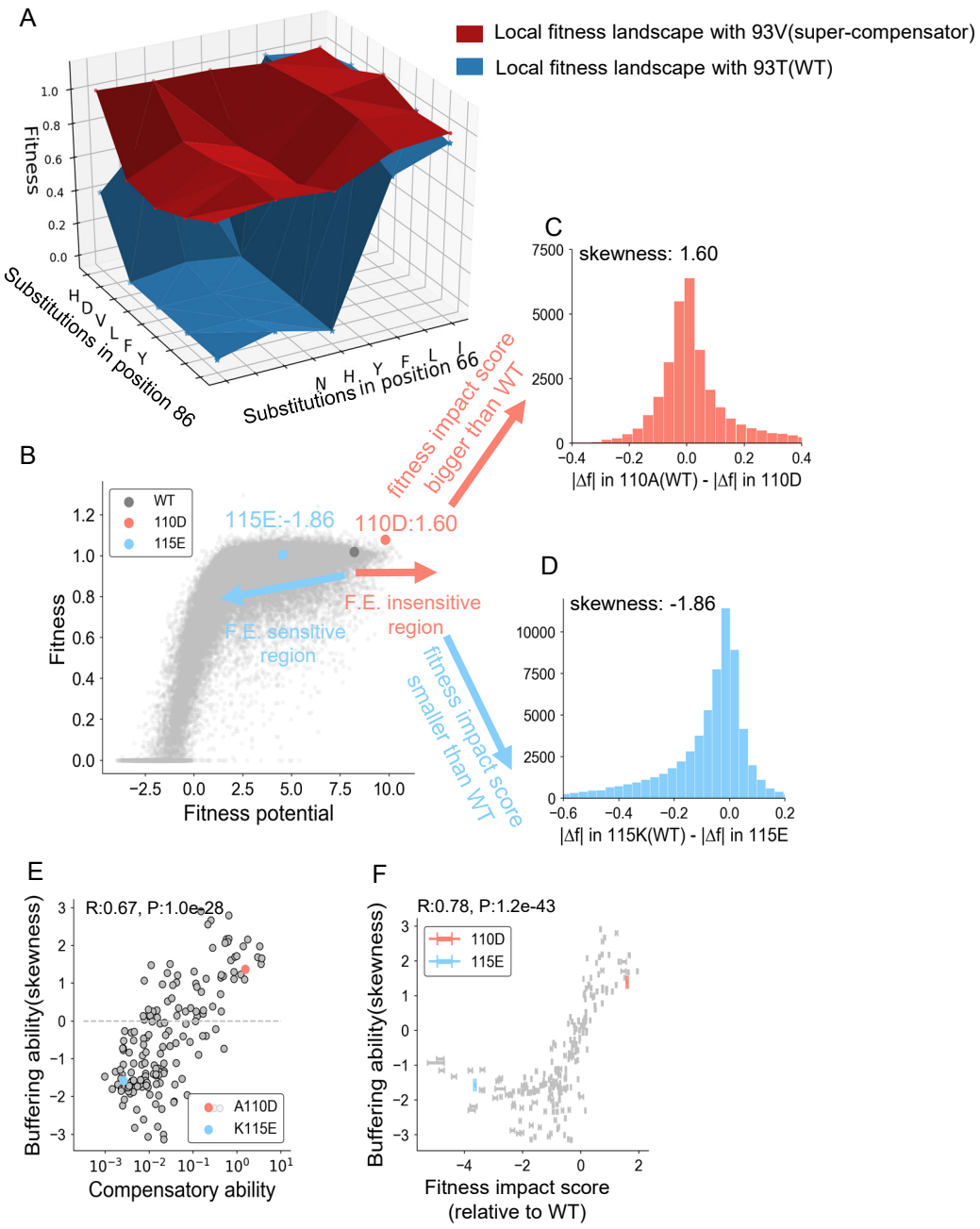

A
Local fitness landscape with 93V(super-compensator)
Local fitness landscape with 93T(WT)
Fitness
Substitutions in position 86
Substitutions in position 66
B
WT
110D
115E
115E:-1.86
110D:1.60
F.E. insensitive region
F.E. sensitive region
Fitness
Fitness potential
fitness impact score bigger than WT
fitness impact score smaller than WT
C
skewness: 1.60
|Δf| in 110A(WT) - |Δf| in 110D
D
skewness: -1.86
|Δf| in 115K(WT) - |Δf| in 115E
E
R:0.67, P:1.0e-28
A110D
K115E
Buffering ability(skewness)
Compensatory ability
F
R:0.78, P:1.2e-43
110D
115E
Buffering ability(skewness)
Fitness impact score
(relative to WT)

Figure 9

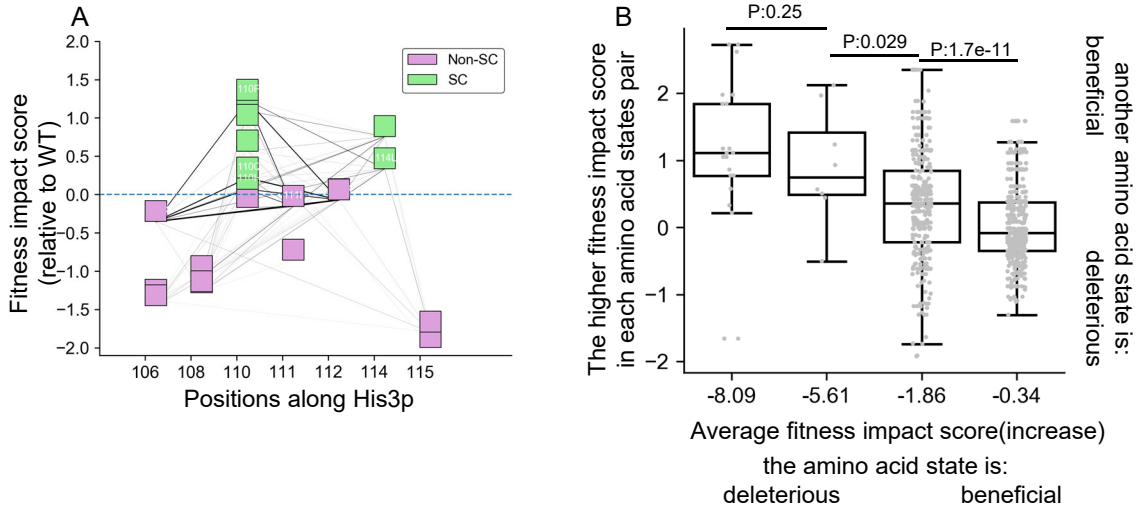

A
Fitness impact score
(relative to WT)
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Non-SC
SC
110F
110K
110E
111I
114L
106
108
110
111
112
114
115
Positions along His3p
B
P:0.25
P:0.029
P:1.7e-11
The higher fitness impact score
in each amino acid states pair
2
1
0
−1
−2
-8.09
-5.61
-1.86
-0.34
Average fitness impact score(increase)
the amino acid state is:
deleterious
beneficial
another amino acid state is:
beneficial
deleterious

Figure 10

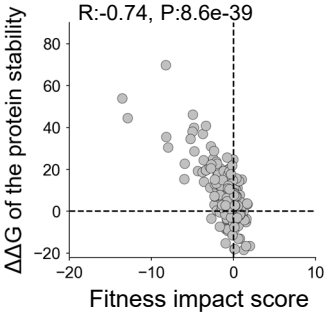

R:-0.74, P:8.6e-39
ΔΔG of the protein stability
80
60
40
20
0
−20
−20
−10
0
10
Fitness impact score

Figure 11

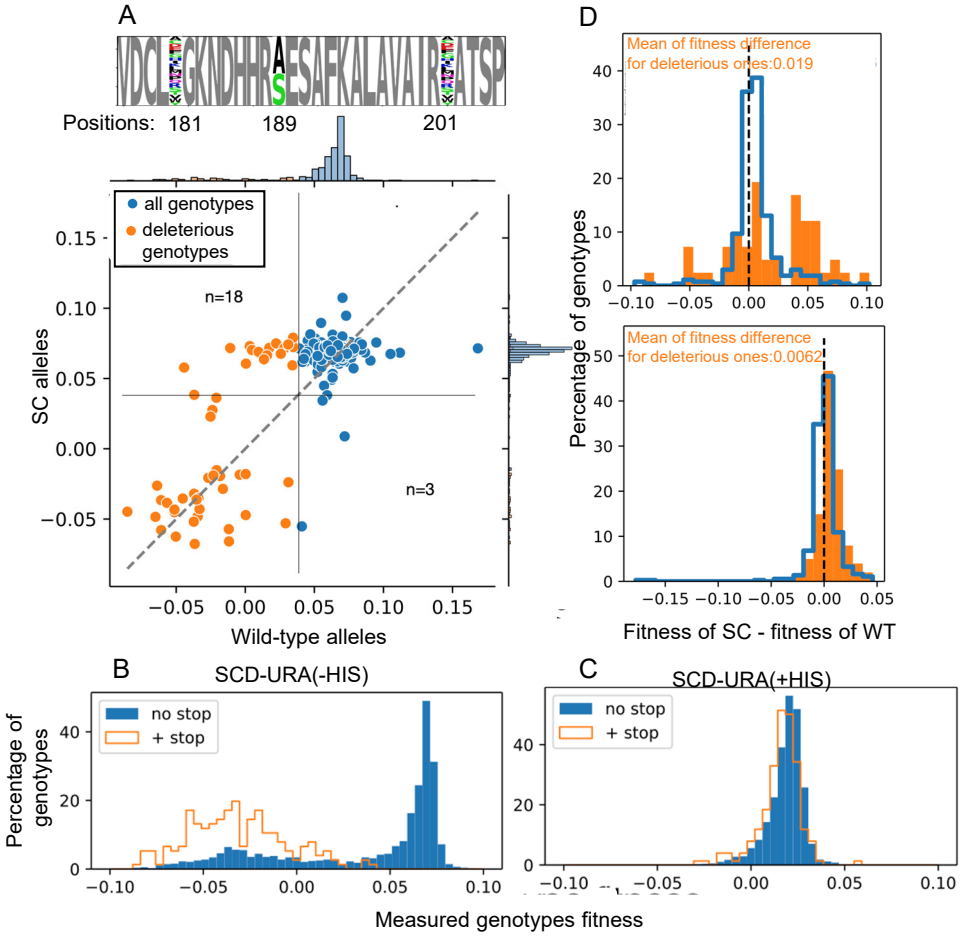

A
Positions: 181
189
201
all genotypes
deleterious genotypes
n=18
n=3
SC alleles
Wild-type alleles
D
Mean of fitness difference for deleterious ones:0.019
Mean of fitness difference for deleterious ones:0.0062
Percentage of genotypes
Fitness of SC - fitness of WT
B
SCD-URA(-HIS)
no stop
+ stop
Percentage of genotypes
C
SCD-URA(+HIS)
no stop
+ stop
Measured genotypes fitness

Figure S1

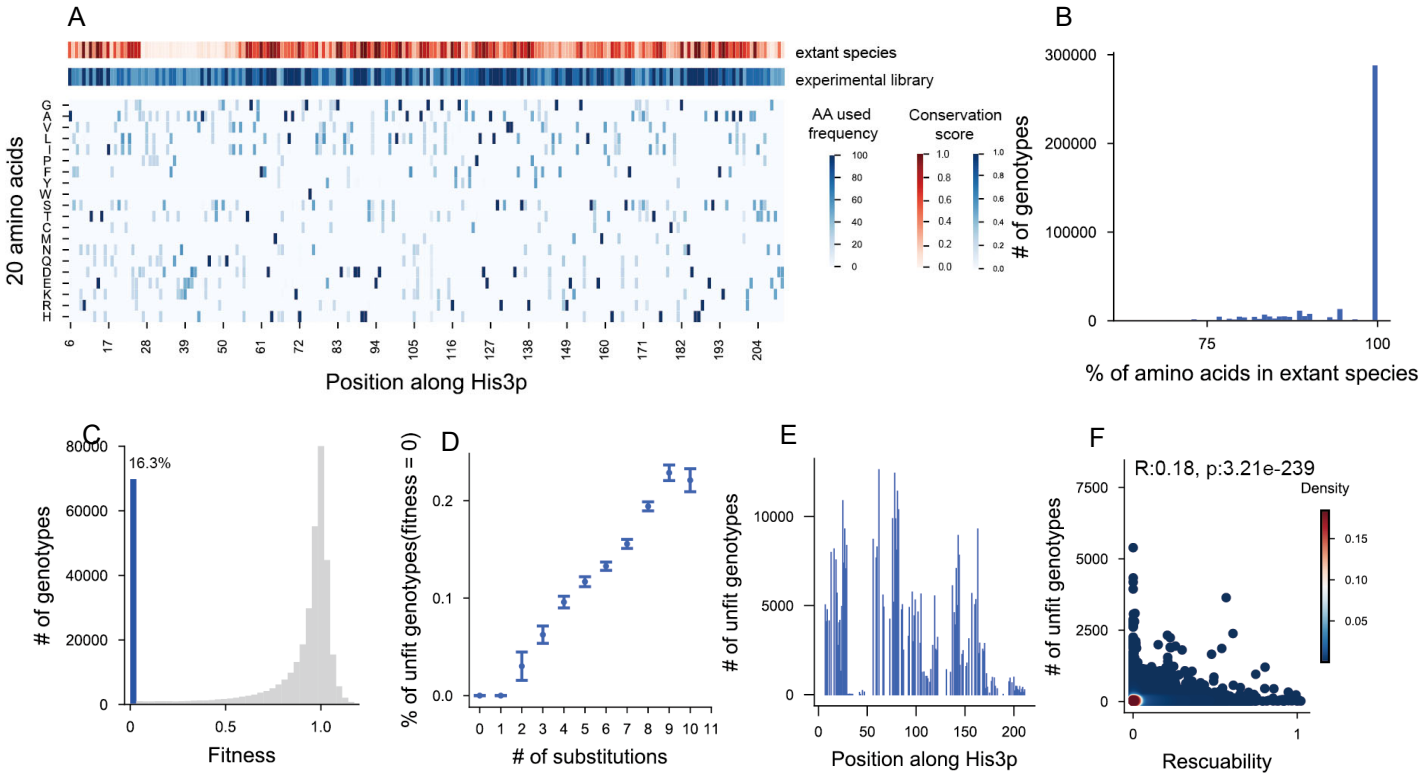

A
extant species
experimental library
20 amino acids
AA used frequency
Conservation score
Position along His3p
B
# of genotypes
% of amino acids in extant species
C
16.3%
# of genotypes
Fitness
D
% of unfit genotypes(fitness = 0)
# of substitutions
E
# of unfit genotypes
Position along His3p
F
R:0.18, p:3.21e-239
Density
# of unfit genotypes
Rescuability

Figure S2

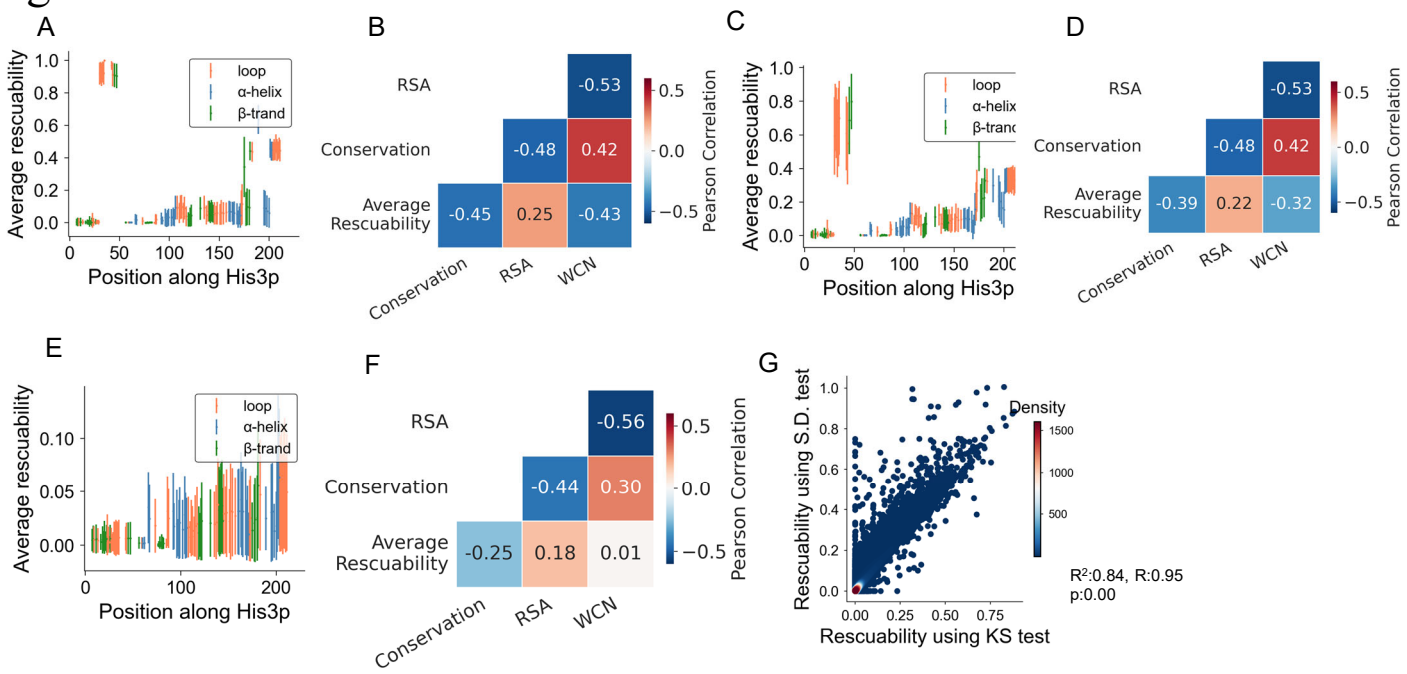

A
Average rescuability
Position along His3p
loop
α-helix
β-trand
B
RSA
Conservation
Average Rescuability
-0.53
-0.48
0.42
-0.45
0.25
-0.43
Conservation
RSA
WCN
Pearson Correlation
C
Average rescuability
Position along His3p
loop
α-helix
β-trand
D
RSA
Conservation
Average Rescuability
-0.53
-0.48
0.42
-0.39
0.22
-0.32
Conservation
RSA
WCN
Pearson Correlation
E
Average rescuability
Position along His3p
loop
α-helix
β-trand
F
RSA
Conservation
Average Rescuability
-0.56
-0.44
0.30
-0.25
0.18
0.01
Conservation
RSA
WCN
Pearson Correlation
G
Rescuability using S.D. test
Rescuability using KS test
Density
R²:0.84, R:0.95
p:0.00

Figure S3

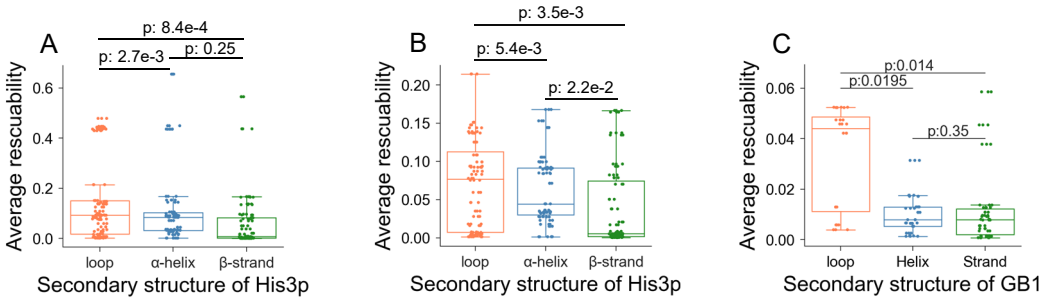

A
p: 8.4e-4
p: 2.7e-3
p: 0.25
Average rescuability
0.0
0.2
0.4
0.6
loop
α-helix
β-strand
Secondary structure of His3p
B
p: 3.5e-3
p: 5.4e-3
p: 2.2e-2
Average rescuability
0.00
0.05
0.10
0.15
0.20
loop
α-helix
β-strand
Secondary structure of His3p
C
p:0.014
p:0.0195
p:0.35
Average rescuability
0.00
0.02
0.04
0.06
loop
Helix
Strand
Secondary structure of GB1

Figure S4

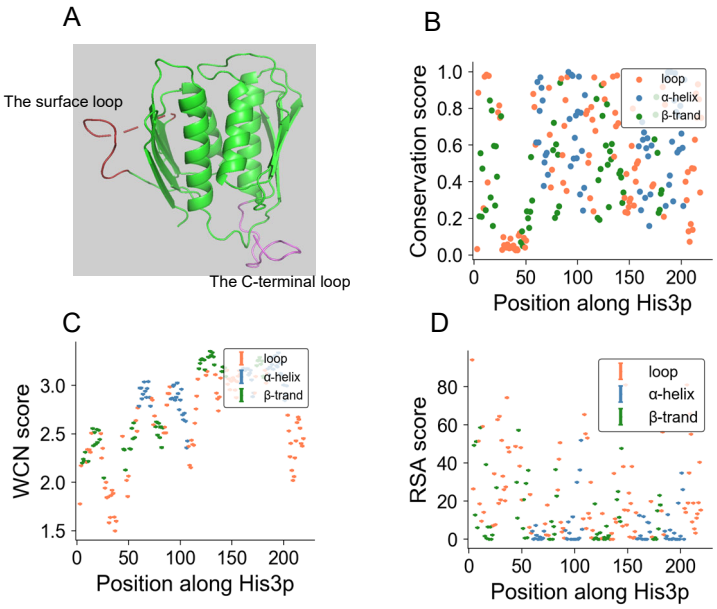

A
The surface loop
The C-terminal loop
B
Conservation score
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
Position along His3p
loop
α-helix
β-trand
C
WCN score
3.0
2.5
2.0
1.5
0
50
100
150
200
Position along His3p
loop
α-helix
β-trand
D
RSA score
80
60
40
20
0
0
50
100
150
200
Position along His3p
loop
α-helix
β-trand

Figure S5

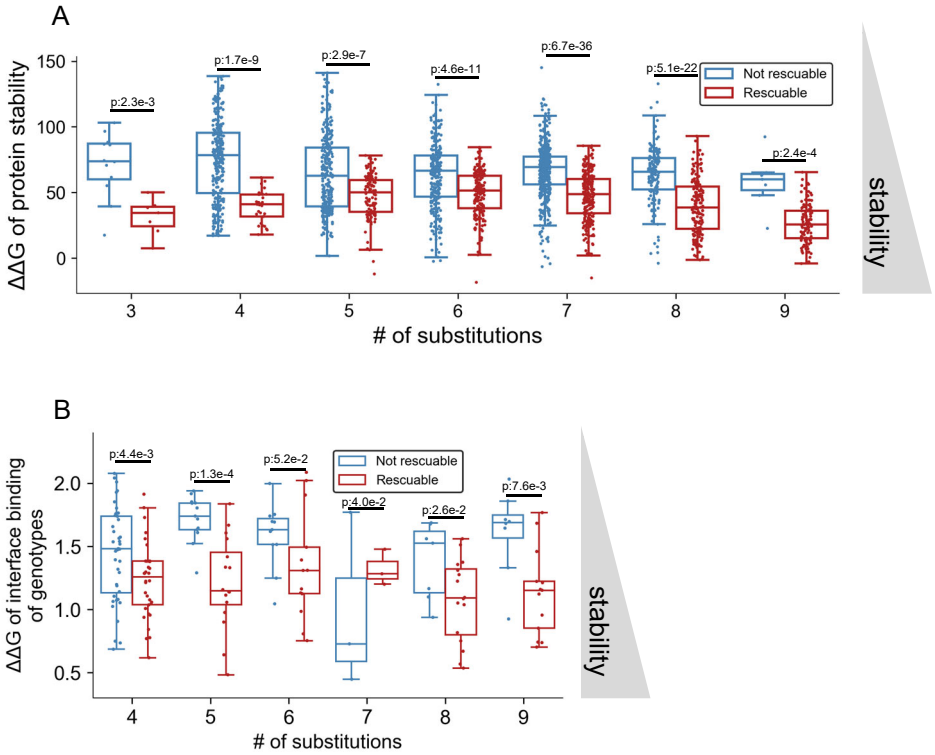

A
ΔΔG of protein stability
150
100
50
0
p:2.3e-3
p:1.7e-9
p:2.9e-7
p:4.6e-11
p:6.7e-36
p:5.1e-22
p:2.4e-4
Not rescuable
Rescuable
3
4
5
6
7
8
9
# of substitutions
stability
B
ΔΔG of interface binding of genotypes
2.0
1.5
1.0
0.5
p:4.4e-3
p:1.3e-4
p:5.2e-2
p:4.0e-2
p:2.6e-2
p:7.6e-3
Not rescuable
Rescuable
4
5
6
7
8
9
# of substitutions
stability

Figure S6

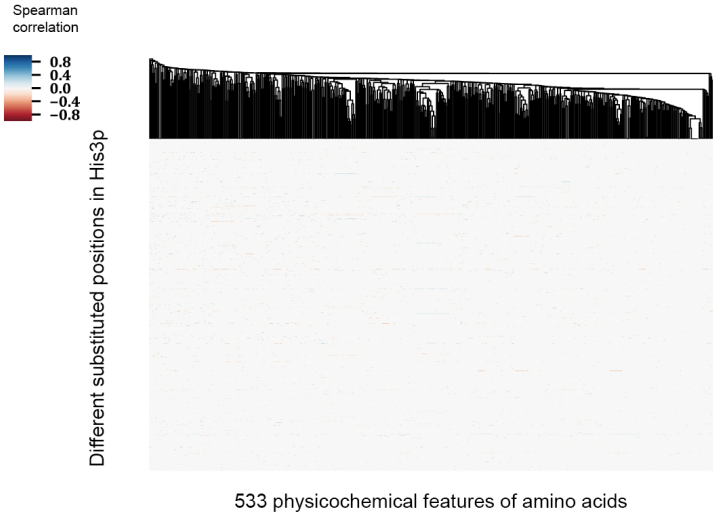

Spearman correlation
0.8
0.4
0.0
−0.4
−0.8
Different substituted positions in His3p
533 physicochemical features of amino acids

Figure S7

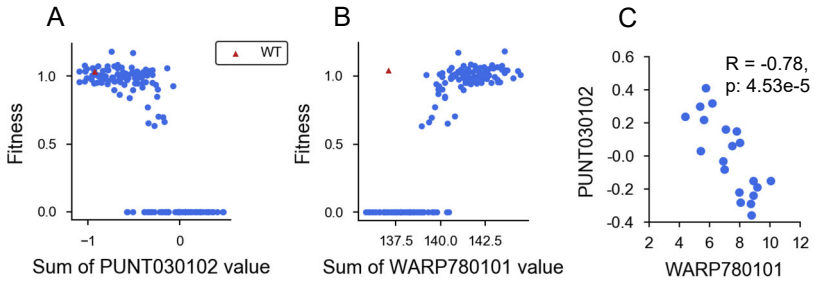

A
WT
Fitness
1.0
0.5
0.0
−1
0
Sum of PUNT030102 value
B
Fitness
1.0
0.5
0.0
137.5
140.0
142.5
Sum of WARP780101 value
C
R = -0.78,
p: 4.53e-5
PUNT030102
0.6
0.4
0.2
-0.0
-0.2
-0.4
2
4
6
8
10
12
WARP780101

Figure S8

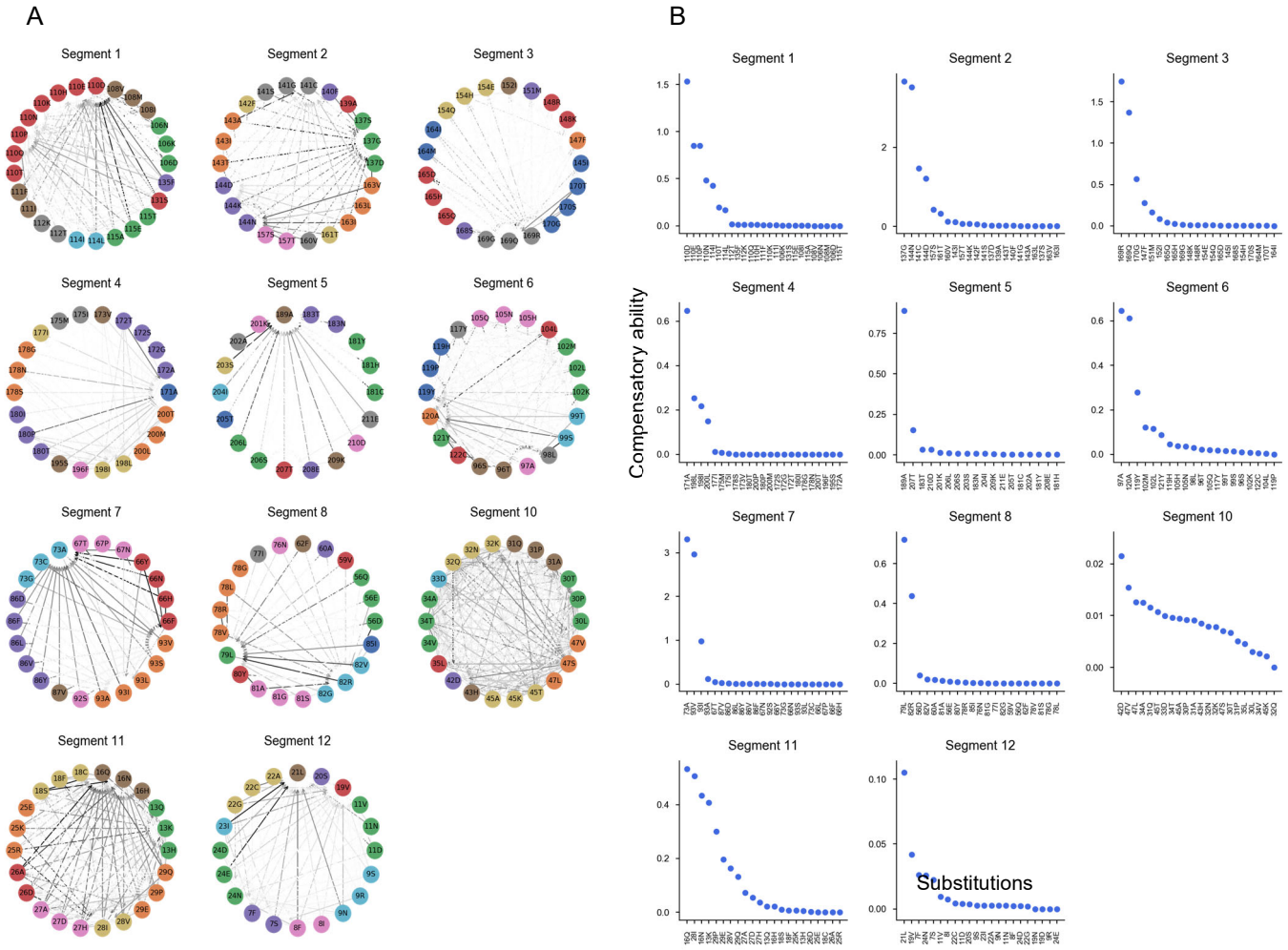

A
Segment 1
Segment 2
Segment 3
Segment 4
Segment 5
Segment 6
Segment 7
Segment 8
Segment 10
Segment 11
Segment 12
B
Compensatory ability
Substitutions

Figure S9

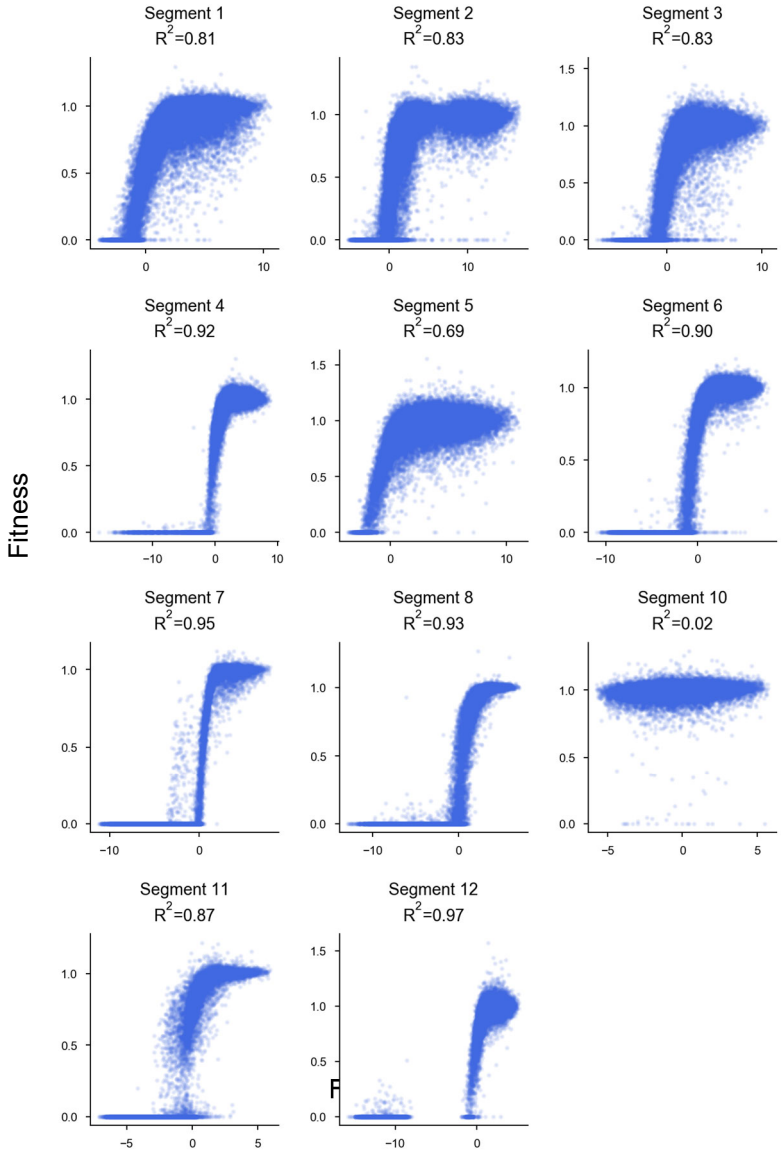

Segment 1
R2=0.81
Segment 2
R2=0.83
Segment 3
R2=0.83
Segment 4
R2=0.92
Segment 5
R2=0.69
Segment 6
R2=0.90
Segment 7
R2=0.95
Segment 8
R2=0.93
Segment 10
R2=0.02
Segment 11
R2=0.87
Segment 12
R2=0.97
Fitness

Figure S10

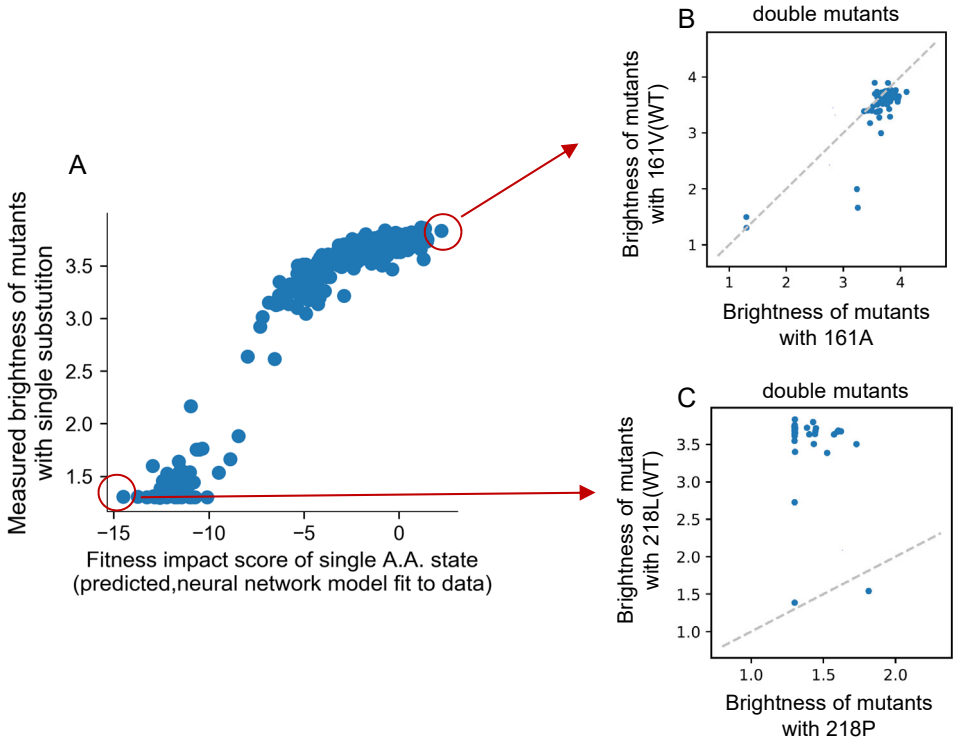

A
Measured brightness of mutants with single substutiton
3.5
3.0
2.5
2.0
1.5
−15
−10
−5
0
Fitness impact score of single A.A. state (predicted,neural network model fit to data)
B
double mutants
Brightness of mutants with 161V(WT)
4
3
2
1
1
2
3
4
Brightness of mutants with 161A
C
double mutants
Brightness of mutants with 218L(WT)
3.5
3.0
2.5
2.0
1.5
1.0
1.0
1.5
2.0
Brightness of mutants with 218P

Figure S11

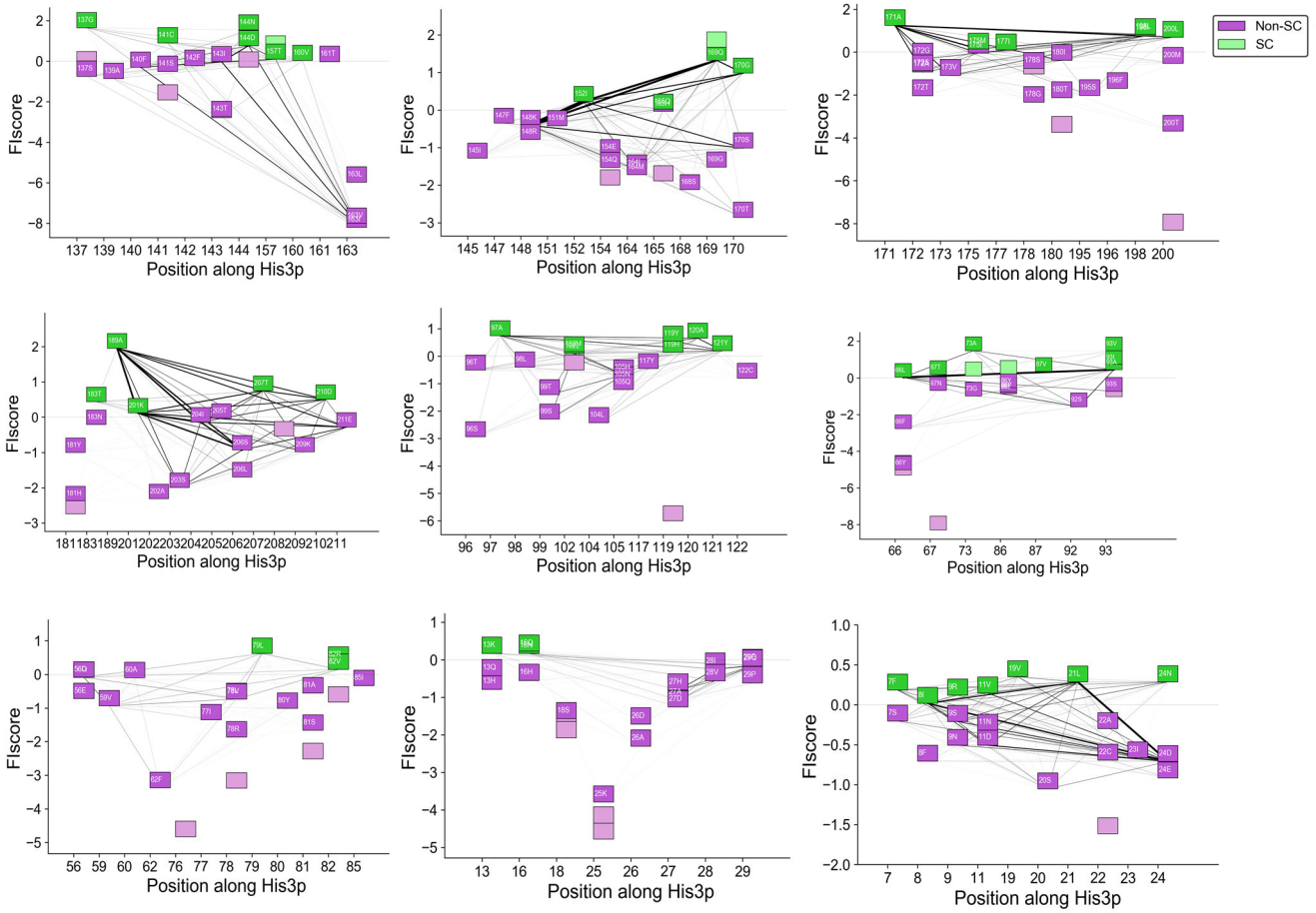
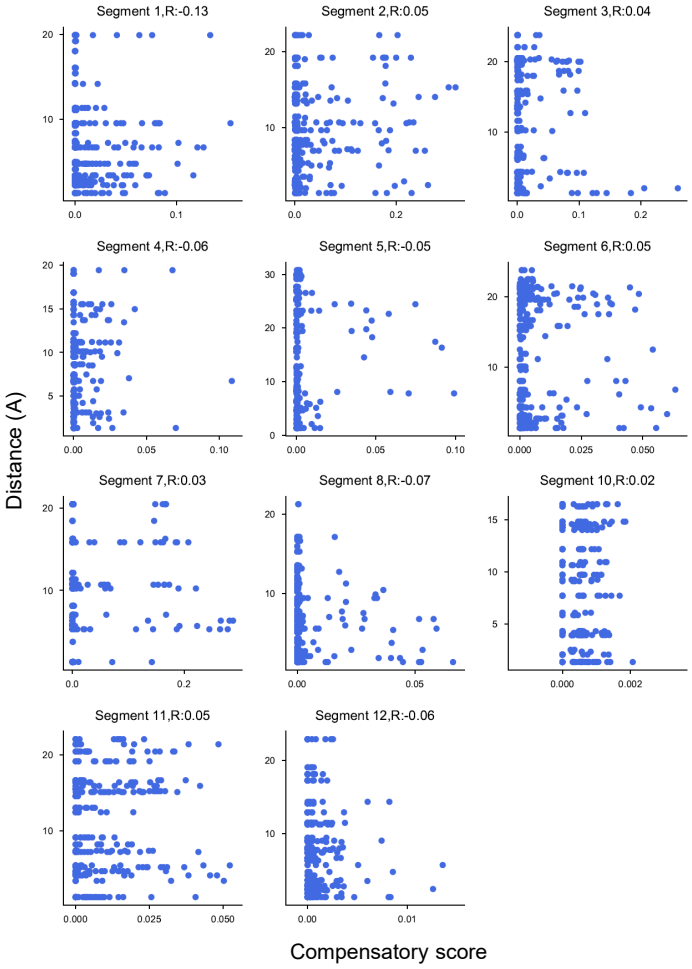


Figure S12


Segment 1,R:-0.13
Segment 2,R:0.05
Segment 3,R:0.04
Segment 4,R:-0.06
Segment 5,R:-0.05
Segment 6,R:0.05
Segment 7,R:0.03
Segment 8,R:-0.07
Segment 10,R:0.02
Segment 11,R:0.05
Segment 12,R:-0.06
Distance (A)
Compensatory score