`

Investigating the influence of environment on the evolution of Hsp90 using comprehensive fitness maps

Julia M. Flynn[1], Ammeret Rossouw[1], Pamela A. Cote-Hammarlof[1], Ines Fragata[2], David Mavor[1], Carl Hollins III[1], Claudia Bank[2], and Daniel N.A. Bolon[1,*]

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01655, USA

[2]Instituto Gulbenkian de Ciência, 2780-156 Oeiras, Portugal

[*]Correspondence: dan.bolon@umassmed.edu

SUMMARY

Gene-environment interactions have long been theorized to influence molecular evolution. However, the environmental dependence of most mutations remains unknown. Using deep mutational scanning, we engineered budding yeast with all 44,604 single codon changes encoding 14,160 amino acid variants in Hsp90 and quantified growth effects under standard laboratory conditions and under five stress conditions (elevated temperature, nitrogen starvation, elevated salinity, high ethanol concentration, and oxidative stress caused by diamide). To our knowledge these are the largest comprehensive fitness maps of point mutant growth effects that have been determined. The growth effects of many variants differed between each of the conditions, indicating that environmental conditions can have a large impact on the evolution of Hsp90. Multiple variants provided growth advantages relative to wildtype Hsp90 under individual conditions, however these variants tended to exhibit growth defects in other environments. The diversity of Hsp90 sequences observed in extant eukaryotes preferentially contain amino acid variants that supported robust growth under all tested conditions. Thus, rather than favoring substitutions in individual conditions, the long-term selective pressure on Hsp90 may have been that of fluctuating environments, leading to robustness under a variety of conditions.

INTRODUCTION

The role of environment has been contemplated in theories of evolution for over a hundred years [1-3], yet molecular level analyses of how environment impacts the evolution of gene sequences remain experimentally under-explored. Depending on environmental conditions, mutations can be categorized into three classes: deleterious mutations that are purged from populations by purifying selection, nearly-neutral mutations that are governed by stochastic processes, and beneficial mutations that provide a selective advantage [4]. It has long been clear that environmental conditions can alter the fitness effects of mutations [5]. However, examining how environmental conditions impact any of the three classes of mutations is challenging. Measurable properties of nearly-neutral and deleterious mutations in natural populations are impacted by both demographics and selection [4], which are difficult to disentangle. In addition, many traits are complex, making it challenging to identify all of the contributing genetic variations [6]. For these reasons, we do not have a detailed understanding of how environmental conditions impact the evolution of most gene sequences.

Analyses of adaptive mutations in both lab experiments and natural populations indicate that they are frequently rare compared to nearly-neutral and deleterious mutations [4,7,8]. Approaches have been developed to identify adaptive mutations in natural populations [9] and lab experiments [10,11]. Most laboratory studies of adaptation rely on stochastic sampling of mutations that identify a subset of all possible adaptive mutations. With stochastic approaches, many potential adaptive mutations can go undetected. Nevertheless, studies sparked by correlating a genotype and a specific environmental condition have revealed a handful of striking evolutionary mechanisms. In one of the earliest studies of molecular adaptation, the correlation of the prevalence of the sickle cell gene in regions where malaria was endemic led to the realization that this gene provided protection from malaria infection [12]. Similar approaches have been utilized to understand other mechanisms of adaptation including how distinct land environments select for coat color of mice mediated by the Agouti gene [13]. Making the link between genotype, phenotype, and adaptation usually involves a tremendous amount of field work, making it difficult to pursue.

On the other end of the spectrum, the environmental dependence of deleterious mutations has been challenging to analyze because of the large number of mutations in this class. The environmental dependence of gene knockouts was investigated in budding yeast [14]. This study revealed that while 19% of genes were essential in rich media, the majority of genes exhibited growth effects under altered conditions. While this result emphasizes that environmental conditions can dramatically influence selection pressure on genes, it does not provide insights into the impact of environmental conditions on gene sequences.

Modern mutational scanning approaches [15] provide novel opportunities to examine fitness effects of the same mutations under different laboratory conditions [16,17]. The EMPIRIC (Exceedingly Meticulous and Parallel Investigation of Randomized Individual Codons) approach

that we developed is particularly well suited to address questions regarding the environmental impact of mutational effects for three reasons: it quantifies growth rates that are a direct measure of experimental fitness, all point mutations are engineered providing comprehensive maps of growth effects, and all the variants can be tracked in the same flask while experiencing identical growth conditions [18]. We have previously used the EMPIRIC approach to investigate how protein fitness maps of ubiquitin vary in different environmental conditions [19]. The analysis of ubiquitin fitness maps revealed that stress environments can exacerbate the fitness defects of mutations. However, the small size of ubiquitin and the near absence of natural variation in ubiquitin sequences (only three amino acid differences between yeast and human) hindered investigation of the properties underlying historically observed substitutions.

Mutational scanning approaches have emerged as a robust method to analyze relationships between gene sequence and function, including aspects of environment-dependent selection pressure. Multiple studies have investigated resistance mutations that enhance growth in drug or antibody environments [20-24]. Most of these studies have focused on interpreting adaptation in the light of protein structure. Of note, Dandage, Chakraborty and colleagues explored how environmental perturbations to protein folding influenced tolerance of mutations in the 178 amino acid gentamicin-resistant gene in bacteria [25]. However, the question of how environmental variation shapes the selection pressure on gene sequences has not been well studied.

Here, we report comprehensive experimental fitness maps of Heat Shock Protein 90 (Hsp90) under multiple stress conditions and compare our experimental results with the historical record of hundreds of Hsp90 substitutions accrued during its billion years of evolution in eukaryotes. Hsp90 encodes a 709 amino acid protein and to our knowledge it is the largest gene for which a comprehensive protein fitness map has been determined. Hsp90 is an essential and highly abundant molecular chaperone which is induced by a wide variety of environmental stresses [26,27]. Hsp90 assists cells in responding to these stressful conditions by facilitating the folding and activation of client proteins through a series of ATP-dependent conformational changes mediated by co-chaperones [28]. These clients are primarily signal transduction proteins, highly enriched in kinases and transcription factors [29] and thus, Hsp90 activity is linked to virtually every process in the cell and influences all facets of cellular physiology. Because Hsp90 helps cells to cope with stress, we hypothesized that the functional pressures acting on Hsp90 may vary in different stress conditions.

The conditions in natural environments often fluctuate, and all organisms contain stress response systems that aid in the acclimation to new conditions. The conditions experienced by different populations and species can vary tremendously depending on the niches that they inhabit, providing the potential for distinct selective pressures on Hsp90. Previous studies of a nine amino acid loop in Hsp90 identified multiple amino acid changes that increased the growth rate of yeast in elevated salinity [30], demonstrating the potential for beneficial mutations in

Hsp90. However, the sequence of Hsp90 is strongly conserved in eukaryotes (57% amino acid identity from yeast to human), indicating consistent strong purifying selection.

To investigate the potential influence of the environment on Hsp90 evolution, we quantified fitness maps in six different conditions. The different conditions impose distinct molecular constraints on Hsp90 sequence. While proximity to ATP is the dominant functional constraint in standard conditions, the influence of client and co-chaperone interactions on growth rate dramatically increases under stress conditions. Increased selection pressure from heat and diamide stresses led to a greater number of beneficial variants compared to standard conditions. The observed beneficial variants were enriched at functional hotspots in Hsp90. However, the natural variants of Hsp90 tend to be robust to all environments tested, indicating selection for robustness to diverse stress conditions in the natural evolution of Hsp90.

RESULTS

We developed a powerful experimental system to analyze the growth rate supported by all possible Hsp90 point mutations under varied growth conditions. Bulk competitions of yeast with a deep sequencing readout enabled the simultaneous quantification of 98% of possible amino acid changes (Figure 1A). The single point mutant library was engineered by incorporating a single degenerate codon (NNN) into an otherwise wildtype Hsp90 sequence as previously described [31]. To provide a sensitive readout of changes in Hsp90 function, we used a plasmid system that reduced Hsp90 protein levels to near-critical levels [32]. We employed a barcoding approach to efficiently track all variants in a single competition flask so that all variants experience identical conditions. As described in the Methods, the barcode strategy enabled us to track mutations across a large gene using a short sequencing readout. The barcode strategy also reduced the impact of misreads as they result in unused barcodes that were discarded from analyses.

We transformed the plasmid library of comprehensive Hsp90 point mutations into a conditional yeast strain where we could turn selection of the library on or off. We used a yeast Hsp90-shutoff strain in which expression of the only genomic copy of Hsp90 is strictly regulated with a galactose-inducible promoter [32]. Yeast containing the mutant libraries were amplified under conditions that select for the plasmid, but not for the function of Hsp90 variants. We switched the yeast to dextrose media to shut off the expression of wildtype Hsp90 and then split the culture into six different environmental conditions. We extracted samples from each condition at multiple time points and used Illumina sequencing to estimate the frequency of each Hsp90 variant over time. We assessed the selection coefficient of each Hsp90 variant from the change in frequency relative to wildtype Hsp90 using a previously developed Bayesian MCMC method [33,34].

To analyze reproducibility of the growth competition, we performed a technical replicate under standard conditions. We used a batch of the same transformed cells that we had frozen and stored such that the repeat bulk competition experiments and sequencing were performed independently. Selection coefficients between replicates were strongly correlated (R^2=0.90), and indicated that we could clearly distinguish wildtype-like mutants from highly deleterious stop-like mutants (Figure 1B). The selection coefficients in this study also correlated strongly (R^2=0.87) with estimates of the Hsp90 N-domain in a previous study [35] (Figure S1A), indicating that biological replicates also show high reproducibility. Of note, variants with strongly deleterious effects exhibited the greatest variation between replicates, consistent with the noise inherent in estimating the frequency of rapidly depleting variants.

The large number of signaling pathways that depend on Hsp90 [29] and its strong sequence conservation suggest that Hsp90 may be sensitive to mutation. However, most variants of Hsp90 were experimentally tolerated in standard conditions (Figure 1C). All possible mutations were compatible with function at 425 positions. Only 18 positions had low mutational tolerance to the extent that 15 or more substitutions caused null-like growth defects (R32, E33, N37, D40, D79, G81, G94, I96, A97, S99, G118, G121, G123, Y125, F156, W300, and R380). All of these positions except for W300 are in contact with ATP or mediate ATP-dependent conformational changes in the N-domain of Hsp90. In fact, the average selection coefficient at different positions (a measure of mutational sensitivity) in standard growth conditions correlates (R^2=0.49) with distance from ATP (Figure S1B). While W300 does not contact ATP, it transmits information from client binding to long range conformational changes of Hsp90 that are driven by ATP hydrolysis [36]. Our results indicate that ATP binding and the conformational changes driven by ATP hydrolysis impose dominant physical constraints throughout the entirety of Hsp90 under standard laboratory conditions.

At first sight, the observation that most mutations were compatible with robust growth in standard conditions is at odds with the fact that the Hsp90 sequence is strongly conserved across large evolutionary distances (Figure 1C). One potential reason for this discrepancy could be that the strength of purifying selection in large natural populations over long evolutionary time-scales is more stringent than can be measured in the laboratory. In other words, experimentally unmeasurable fitness defects could be subject to purifying selection in nature. In addition, the range of environmental conditions that yeast experience in natural settings may not be reflected by standard laboratory growth conditions. To investigate the impact of environmental conditions on mutational effects in Hsp90, we measured the growth rate of Hsp90 variants under five additional stress conditions.

Impact of stress conditions on mutational sensitivity of Hsp90

We measured the fitness of Hsp90 variants in conditions of nitrogen depletion (ND) (0.0125% ammonium sulfate), hyper-osmotic shock (0.8 M NaCl), ethanol stress (7.5% ethanol), the

sulfhydryl-oxidizing agent diamide (0.85 mM), and temperature shock (37°C). All of these stresses are known to elicit a common shared environmental stress response characterized by altered expression of ~900 genes as well as having specific responses unique to each stress [27]. Genes encoding heat shock proteins, including Hsp90, are transiently upregulated in all these stresses except elevated salinity [27,37].

One way to characterize stress conditions is to measure the extent to which they slow down growth. For our experiments, each of the environmental stresses were selected to partially decrease the growth rate. Consistently, all stresses reduced the growth rate of the parental strain within a two-fold range, with depletion of nitrogen levels causing the smallest reduction in growth rate and diamide causing the greatest reduction (Figure 2A). To investigate how critical Hsp90 is for growth in each condition, we measured growth rates of yeast with either normal or more than 10-fold reduced [32] levels of Hsp90 protein (Figure 2A). Under standard conditions, the normal level of Hsp90 protein can be dramatically reduced without major impacts on growth rate, consistent with previous findings [32,38].

We anticipated that Hsp90 would be required at increased levels for robust experimental growth in diamide, nitrogen starvation, ethanol, and high temperature [27] based on the concept that cells increase expression level of genes in conditions where those gene products are needed at higher concentration. Consistent with this concept, reduced Hsp90 levels cause a marked decrease in growth rate at 37°C. However, Hsp90 protein levels had smaller impacts on growth rates under the other stress conditions, indicating that reliance on overall Hsp90 function does not increase dramatically in these conditions.

We quantified the growth rates of all Hsp90 single-mutant variants in each of the stress conditions as selection coefficients where 0 represents wildtype and -1 represents null alleles (Figure S2A-F). We could clearly differentiate between the selection coefficients of wildtype synonyms and stop codons in all conditions (Figure 2B), and we normalized to these classes of mutations to facilitate comparisons between each condition. Of note, the observed selection coefficients of wildtype synonyms varied more in conditions of high temperature and diamide stress compared to standard (Figure S2G). We also note greater variation in the selection coefficients of barcodes for the same codon in the diamide and high temperature conditions (Figure S2H). We conclude that diamide and elevated temperature provided greater noise in our selection coefficient measurements. To take into account differences in signal to noise for each condition, we either averaged over large numbers of mutations or categorized selection coefficients as wildtype like, strongly deleterious, intermediate, or beneficial based on the distribution of wildtype synonyms and stop codons in each condition (see Materials and Methods and Figure S2I).

We compared selection coefficients of each Hsp90 variant in each stress condition to standard condition (Figure 2B&C). The stresses of 37°C and diamide tend to exaggerate the growth defects of many mutants compared to standard conditions, whereas high salt and ethanol tend to rescue growth defects (Figure 2B&C and S2J). According to the theory of metabolic flux

[39,40], gene products that are rate limiting for growth will be subject to the strongest selection. Accordingly, the relationship between Hsp90 function and growth rate should largely determine the strength of selection acting on Hsp90 sequence. Conditions where Hsp90 function is more directly linked to growth rate would be more sensitive to Hsp90 mutations than conditions where Hsp90 function can be reduced without changing growth rates [32,41]. The average selection coefficients are more deleterious in diamide and temperature stress compared to standard conditions. These findings are consistent with heat and diamide stresses causing a growth limiting increase in unfolded Hsp90 clients that is rate limiting for growth. In contrast, the average selection coefficients are less deleterious in ethanol and salt stress than in standard conditions, consistent with a decrease in the demand for Hsp90 function in these conditions. Of note, the biophysical impacts of these sets of conditions are distinct: increased temperature and diamide both generally act to increase protein unfolding, while salt and ethanol have more complex impacts that include inducing aggregation.

Structural analyses of environmental responsive positions

Altering environmental conditions had a pervasive influence on mutational effects along the sequence of Hsp90 (Figure 3A & S3A). We structurally mapped the average selection coefficient of each position in each condition relative to standard conditions as a measure of the sensitivity to mutation of each position under each environmental stress (Figure 3A). Many positions had mutational profiles that were responsive to a range of environments. Environmentally responsive positions with large changes in average selection coefficient in at least three conditions are highlighted on the Hsp90 structure in green in Figure 3B. Unlike the critical positions that cluster around the ATP binding site (Figure 1C), the environmentally responsive positions are located throughout all domains of Hsp90. Similar to critical residues, environmentally responsive positions are more conserved in nature compared to other positions in Hsp90 (Figure 3C), suggesting that the suite of experimental stress conditions tested captured aspects of natural selection pressures on Hsp90 sequence.

Hsp90 positions with environmentally responsive selection coefficients were enriched in binding contacts with clients, co-chaperones and intramolecular Hsp90 contacts involved in transient conformational changes (Figure 3D and S3B). About 65% of the environmentally responsive residues have been identified either structurally or genetically as interacting with binding partners [42-55], compared to about 15% of positions that were not responsive to stress conditions. While ATP binding and hydrolysis are the main structural determinants that constrain fitness in standard growth conditions, client and co-chaperone interactions have a larger impact on experimental fitness under stress conditions. Although the mean selection coefficients of mutations at the known client and co-chaperone binding sites are responsive to changes in environment, the direction of the shift of growth rate compared to standard conditions depends on the specific binding partner and environment (Figure 3E). This suggests that different environments place unique functional demands on Hsp90 that may be mediated

by the relative affinities of different binding partners. Consistent with these observations, we hypothesize that Hsp90 client priority is determined by relative binding affinity and that Hsp90 mutations can reprioritize clients that in turn impacts many signaling pathways.

Constraint of mutational sensitivity at high temperature

We find that different environmental conditions impose unique constraints on Hsp90, with elevated temperature placing the greatest purifying selection pressure on Hsp90. Of the 2504 variants of Hsp90 that are deleterious when grown at 37°C, 884 of them (~35%) are deleterious only in this condition (Figure 3F). We defined mutants that confer temperature sensitive (*ts*) growth phenotypes on cells as variants with selection coefficients within the distribution of wildtype synonyms in standard conditions and that of stop codons at 37°C. Based on this definition, 675 Hsp90 amino acid changes (roughly 5% of possible changes) were found to be temperature sensitive (Figure 4A). We sought to understand the physical underpinnings of this large set of Hsp90 *ts* mutations.

We examined Hsp90 *ts* mutations for structural and physical patterns. We found that *ts* mutations tended to concentrate at hotspots (Figure 4B). These hotspots were spread across all three domains of Hsp90 (Figure 4C). The largest cluster of hotspots occurred in the C domain of Hsp90. The C domain forms a constitutive homodimer that is critical for function [56]. Of note, homo-oligomerization domains may have a larger *ts* potential because all subunits contribute to folding and dimerization essentially multiplying the impacts of mutations [57]. To explore the physical underpinnings of *ts* mutations we examined if they were buried in the structure or surface exposed. Mutations at buried residues tend to have a larger impact on protein folding energy compared to surface residues [58]. Consistent with the idea that many *ts* mutations may disrupt protein folding at elevated temperature, substitutions that confer a *ts* phenotype are enriched in buried residues (Figure 4D). Also consistent with this idea, *ts* mutations tend to have negative Blosum scores (Figure 4E), a hallmark of disruptive amino acid changes.

Because growth at elevated temperatures requires higher levels of Hsp90 protein [59], some *ts* mutations may be due to a reduced function that is enough for growth at standard temperature, but is insufficient at 37 °C [54]. We reasoned that we could distinguish these mutants by examining how growth rate depended on the expression levels of Hsp90. We expect that destabilizing mutants that cause Hsp90 to unfold at elevated temperature would not support efficient growth at 37°C independent of expression levels. In contrast, we expect mutants that reduce Hsp90 function to exhibit an expression-dependent growth defect at 37 °C. We tested a panel of *ts* mutations identified in the bulk competitions at high and low expression levels (Figure 4F). The dependence of growth rate at 37 °C on expression level varied for different Hsp90 *ts* variants. The I64D, I66E and L499R Hsp90 mutants have no activity at 37°C irrespective of expression levels. These disruptive substitutions at buried positions likely destabilize the structure of Hsp90. In contrast, increasing the Hsp90 expression levels at least

partially rescued the growth defect for five *ts* variants (L50D, K102A, D180L, K398L, K594I), indicating that these variants do not providing enough Hsp90 function for robust growth at elevated temperature. All five of these expression dependent *ts* variants were located at surface positions, indicating that the location of *ts* mutations can delineate different mechanistic classes.

Hsp90 potential for adaptation to environmental stress

Numerous Hsp90 variants provided a growth benefit compared to the wildtype sequence in stress conditions. The largest number of beneficial variants in Hsp90 occurred in high temperature and diamide conditions (Figure 5A). Multiple lines of evidence indicate that these mutants are truly beneficial variants and not simply measurement noise. First, the beneficial amino acids generally exhibited consistent selection coefficients among synonymous variants (Figure S5A). Second, adaptive mutants in diamide and high temperature cluster at certain positions in a significant manner (see below). Finally, we confirmed the increased growth rate at elevated temperature of a panel of variants analyzed in isolation (Figure S5B). Beneficial mutations in elevated temperature and diamide often clustered at specific positions in Hsp90 (Figure 5B), indicating that the wildtype amino acids at these positions are far from optimum for growth in these conditions. In contrast, the apparent beneficial mutations in other conditions did not tend to cluster at specific positions (Figure S5C).

To obtain a more general picture of the potential for adaptation derived from the full fitness distributions, we used Fisher's Geometric model (FGM) [60]. According to FGM, populations evolve in an n-dimensional phenotypic space, through random single step mutations, and any such mutation that brings the population closer to the optimum is considered beneficial. An intuitive hypothesis derived from FGM is that the potential for adaptation in a given environment (that is the availability of beneficial mutations) depends on the distance to the optimum. In order to estimate the distance to the optimum d, we adopted the approach by Martin and Lenormand and fitted a displaced gamma distribution to the neutral and beneficial mutations for each environment [61]. We observed that the yeast populations were furthest from the optimum in elevated temperature and diamide (d=0.072 and 0.05, respectively), followed by nitrogen deprivation (d=0.023), high salinity and ethanol (d=0.021) and standard (d=0.014). This suggests that elevated temperature and diamide have the largest potential for adaptation and is consistent with the observation of the largest proportions of beneficial mutations in these environments. Interestingly, previous results from a 9-amino-acid region in Hsp90 indicated that there was very little potential for adaptation at high temperature (36°C) as compared with high salinity [62]. This apparent contradiction between results from the full Hsp90 sequence and the 582-590 region indicates that a specific region of the protein may be already close to its functional optimum in a specific environment, whereas there is ample opportunity for adaptation when the whole protein sequence is considered.

In diamide and elevated temperature, the clustered beneficial positions were almost entirely located in the ATP-binding domain and the middle domain (Figure 5C), both of which make extensive contacts with clients and co-chaperones [42-46,55]. Beneficial mutations in elevated temperature and diamide conditions were preferentially located on the surface of Hsp90 (Figure 5D) at positions accessible to binding partners. Analyses of available Hsp90 complexes indicate that beneficial positions were disproportionately located at known interfaces with co-chaperones and clients (Figure 5E). Clustered beneficial mutations are consistent with disruptive mechanisms because different amino acid changes can lead to disruptions, whereas a gain of function is usually mediated by specific amino acid changes. Amino acids that are beneficial in diamide and elevated temperature tend to exhibit deleterious effects in standard conditions (Figure 5F), consistent with a cost of adaptation. We conjecture that the clustered beneficial mutations are at positions that mediate the binding affinity of subsets of clients and co-chaperones and that disruptive mutations at these positions can lead to re-prioritization of multiple clients. The priority or efficiency of Hsp90 for sets of clients can in turn impact most aspects of physiology because Hsp90 clients include hundreds of kinases that influence virtually every aspect of cell biology.

In the first ten amino acids of Hsp90, we noted a large variation in the selection coefficients of synonymous mutations at elevated temperature (Figure S5D). These synonymous mutations were only strongly beneficial at high temperature where Hsp90 protein levels are limiting for growth. Analysis of an individual clones confirms that synonymous mutations at the beginning of Hsp90 that were beneficial at high temperature were expressed at higher level in our plasmid system (Figure S5E, S5F). These results are consistent with a large body of research showing that mRNA structure near the beginning of coding regions often impacts translation efficiency [63-65], and that adaptations can be mediated by changes in expression levels [66]. Outside of the first ten amino acids, we did not observe large variation in selection coefficients of synonymous mutations.

Natural selection favors Hsp90 variants that are robust to environment

We next examined how experimental protein fitness maps compared with the diversity of Hsp90 sequences in current eukaryotes. We analyzed Hsp90 diversity in a set of 267 sequences from organisms that broadly span across eukaryotes. We identified 1750 amino acid differences in total that were located at 499 positions in Hsp90. We examined the experimental growth effects of the subset of amino acids that were observed in nature. While the overall distribution of selection coefficients in all conditions was bimodal with peaks around neutral (s=0) and null (s=-1), the natural amino acids were unimodal with a peak centered near neutral (Figure 6A). The vast majority of natural amino acids had wildtype-like fitness in all conditions studied here (Figure 6B&C). Whereas naturally occurring amino acids in Hsp90 were rarely deleterious in any experimental condition, they were similarly likely to provide a growth benefit compared to all possible amino acids (5%). This observation indicates that condition-dependent fitness benefits

are not a major determinant of natural variation in Hsp90 sequences. Instead, our results indicate that natural selection has favored Hsp90 substitutions that are robust to multiple stressful conditions (Figure 6D).

Epistasis may provide a compelling explanation for the naturally occurring amino acids that we observed with deleterious selection coefficients. Analyses of Hsp90 mutations in the context of likely ancestral states has demonstrated a few instances of historical substitutions with fitness effects that depend strongly on the Hsp90 sequence background [67]. Indeed, many of the natural amino acids previously identified with strong epistasis (E7A, V23F, T13N) are in the small set of natural amino acids with deleterious effects in at least one condition. Further analyses of natural variants under diverse environmental conditions will likely provide insights into historical epistasis and will be the focus of future research.

DISCUSSION

In this study, we analyzed the protein-wide distribution of fitness effects of Hsp90 across standard conditions and five stress conditions. Fitness effects under standard conditions indicated that critical functional residues were largely explained by proximity to ATP. In contrast, positions spread throughout the structure of Hsp90 exhibited selection coefficients that were responsive to different conditions. While the main functional constraints were apparent in standard conditions, stress conditions revealed many additional functional constraints as well as beneficial mutations.

We observed distinct structural trends for mutations that provide environment-dependent costs and benefits. Many mutations in Hsp90 caused growth defects at elevated temperature where Hsp90 function is limiting for growth. These temperature sensitive mutations tended to be buried and in the homodimerization domain, consistent with an increased requirement for folding stability at elevated temperatures. In contrast, beneficial mutations tended to be on the surface of Hsp90 and at contact sites with binding partners, suggesting that change-of-function mutations may be predominantly governed by alterations to binding interactions.

Importantly, our results demonstrate that while mutations to Hsp90 can provide a growth advantage in specific environmental conditions, naturally occurring amino acids in Hsp90 tend to support robust growth over multiple stress conditions. The finding of beneficial mutations in Hsp90 in specific conditions suggests that similar long-term stresses in nature can lead to positive selection on Hsp90. Consistent with previous work [26], we found that experimentally beneficial mutations tended to have a fitness cost in alternate conditions (Figure 5F). This indicates that natural environments which fluctuate among different stresses would reduce or eliminate positive selection on Hsp90. Therefore, our results suggest that natural selection on Hsp90 sequence has predominantly been governed by strong purifying selection integrated over multiple stressful conditions. Taken together, these results support the hypothesis that natural populations might experience a so-called "micro-evolutionary fitness seascape" [68], in

which rapidly fluctuating environments result in a distribution of quasi-neutral substitutions over evolutionary time scales.

Acknowledgements

Thanks to Tyler Starr for providing the alignment of Hsp90 sequences used to assess natural variation. This work was supported by grants from the National Institutes of Health (R01-GM112844 to D.N.A.B. and F32-GM119205 to J.M.F)

FIGURE LEGENDS

Figure 1. Approach to determine protein fitness maps of Hsp90. A. Barcoded competition strategy to analyze the growth effects of all single codon variants of Hsp90 in a single bulk culture. B. Measurements of selection coefficients of all amino acid variants are reproducible in replicate growth competitions. Wildtype amino acids are shown in green and stop codons are shown in red. C. Selection coefficients in standard conditions mapped onto a homodimeric structure of Hsp90 [42] and compared to patterns of evolutionary conservation. ATP is shown in yellow.

Figure 2. Relating impacts of environmental stresses on yeast growth rates and selection on Hsp90 sequence. A. Growth rate of yeast with normal and reduced expression of Hsp90 protein in standard and stress conditions. B. Selection coefficients of Hsp90 amino acid variants in stress conditions compared to standard conditions. Wildtype amino acids are shown in green and stop codons are shown in red. Selection coefficients were scaled to null (s=-1) for the average stop codon and neutral (s=0) for the average wildtype. C. Average selection coefficients of mutations, a metric of the strength of selection acting on Hsp90 sequence, in each stress condition relative to standard conditions.

Figure 3. Environmental stresses elicit distinct patterns of selection on Hsp90. A. The average selection coefficient at each position relative to standard conditions was mapped onto Hsp90 structure [42]. B. Structural images indicating the location of positions that are critical for Hsp90 function in all conditions (magenta) and positions that are environmentally responsive (green). Critical residues have mean selection coefficients that are null-like (within the distribution of stop codons) in all environments. Environmentally responsive positions had mean selection coefficients that differed from standard in three or more environmental conditions by an amount greater than one standard deviation of wildtype synonyms. Tolerant residues were not shifted more than this cutoff in any environment. C. For different classes of positions, evolutionary variation was calculated as amino acid entropy at each position in Hsp90 sequences from diverse eukaryotes. D. Fraction of different classes of mutations located at contact sites with binding partners. E. The average selection coefficient in each environment relative to standard at all the Hsp90 positions at each stated interface. F. Venn diagram of deleterious mutations in different environmental conditions [69]. Total number of deleterious mutants in each condition are stated in parentheses.

Figure 4. Abundance and mechanism of temperature sensitive mutations in Hsp90. A. Temperature sensitive amino acid variants were identified that supported wildtype-like growth at standard (30°C) temperature, but were null-like at 37 °C in bulk competitions. B. The identified *ts* variants tend to cluster at specific positions in Hsp90. C. Mapping positions with multiple *ts* variants onto Hsp90 structure. D. Solvent accessible surface area (SASA) of positions with *ts* mutations compared to positions lacking *ts* mutations. E. Amino acid similarity to the

wildtype was estimated as the Blosum score for *ts* and non-*ts* variants. F. Growth rate of a panel of individual Hsp90 *ts* variants analyzed in isolation.

Figure 5. Beneficial variants in diamide and elevated temperature conditions. A. Number of beneficial mutations identified in each condition based on selection coefficients more than two standard deviations greater than wildtype synonyms. B. Distribution of the number of beneficial mutations at the same position in both 37 °C (left) and diamide (right) conditions. C. Location of positions with four or more beneficial mutations. Positions that are unique to diamide or 37 °C are shown in cyan and two shared positions are shown in magenta. D. The solvent accessible surface area of beneficial positions at 37 °C and in diamide compared to all positions. E. The fraction of Hsp90 positions at interfaces that were beneficial in 37 °C and diamide conditions. F. Selection coefficients in standard conditions for beneficial mutations at 37 °C and in diamide compared to wildtype synonyms.

Figure 6. Experimental growth effects of natural amino acid variants of Hsp90. A. The distribution of selection coefficients of natural variants compared to all variants in each environmental condition. B. Across all environments, the fraction of natural variants compared to all variants that were beneficial, wildtype like, or deleterious. C. The fraction of natural variants compared to all variants that were environmentally responsive or tolerant in all environments. Categories were defined as in Figure 3B. D. Landscape model indicating that natural variants of Hsp90 tend to support robust growth under a variety of stress conditions.

MATERIALS AND METHODS

Generating mutant libraries

A library of Hsp90 genes was saturated with single point mutations using oligos containing NNN codons as previously described [31]. The resulting library was pooled into 12 separate 60 amino acid long sub-libraries along the Hsp90 sequence (amino acids 1-60, 61-120 etc.) and combined via Gibson Assembly (NEB) with a linearized p414ADHΔter Hsp90 destination vector. To simplify sequencing steps during bulk competition, each variant of the library was tagged with a unique barcode. For each 60 amino acid sub-library, a pool of DNA constructs containing a randomized 18-bp barcode sequence (N18) was cloned 200 nt downstream from the Hsp90 stop codon via restriction digestion, ligation, and transformation into chemically competent E. coli with the goal of each mutant being represented by 10-20 unique barcodes.

Barcode association of library variants

We added barcodes and associated them with Hsp90 variants essentially as previously described [67]. To associate barcodes with Hsp90 variants, we performed paired-end sequencing of each 60 amino acid sub-library using a primer that reads the N18 barcode in one read and a primer unique to each sub-library that anneals upstream of the region containing mutations. To facilitate efficient Illumina sequencing, we generated PCR products that were less than 1kb in length for sequencing. We created shorter PCR products by generating plasmids with regions removed between the randomized regions and the barcode. To remove regions from the plasmids, we performed restriction digest with two unique enzymes, followed by blunt ending with T4 DNA polymerase (NEB) and plasmid ligation at a low concentration (3 ng/µL) to favor circularization over bimolecular ligations. The resulting DNA was relinearized by restriction digest, and amplified with 11 cycles of PCR to generate products for Illumina sequencing. The resulting PCR products were sequenced using an Illumina MiSeq instrument with asymmetric reads of 50 bases for Read1 (barcode) and 250 bases for Read2 (Hsp90 sequence). After filtering low-quality reads (Phred scores <10), the data was organized by barcode sequence. For each barcode that was read more than three times, we generated a consensus of the Hsp90 sequence that we compared to wildtype to call mutations.

Bulk Growth Competitions

Equal molar quantities of each sub-library were mixed to form a pool of DNA containing the entire Hsp90 library with each codon variant present at similar concentration. The plasmid library was transformed using the lithium acetate procedure into the DBY288 Hsp90 shutoff strain essentially as previously described [32]. Sufficient transformation reactions were performed to attain ~5 million independent yeast transformants representing a 5-fold sampling

for the average barcode and 100-fold sampling for the average codon variant. Following 12 hours of recovery in SRGal (synthetic 1% raffinose and 1% galactose) media, transformed cells were washed five times in SRGal-W (SRGal lacking tryptophan) media to remove extracellular DNA, and grown in SRGal-W media at 30°C for 48 h with repeated dilution to maintain the cells in log phase of growth. This yeast library were was supplemented with 20% glycerol, aliquoted and slowly frozen by placing a microfuge tubes with cells in a room temperature plastic tube holder in a -80°C freezer.

For each competition experiment, an aliquot of the frozen yeast library cells was thawed at 37°C. Viability of the cells was accessed before and after freezing and was determined to be greater than 90% with this slow freeze, quick thaw procedure. Thawed cells were amplified in SRGal-W for 24 hours, and then shifted to shutoff conditions by centrifugation, washing, and resuspension in 300 mL of synthetic dextrose lacking tryptophan (SD-W) for 12 hours at 30°C. At this point, cells were split and transferred to different conditions including SD-W (standard - 2% dextrose, 0.5% ammonium sulphate, 30°C), ND (nitrogen depletion - 2% dextrose, 0.0125% ammonium sulphate, 30°C), NaCl (2% dextrose, 0.5% ammonium sulphate, 0.8 M NaCl, 30°C), EtOH (2% dextrose, 0.5% ammonium sulphate, 7.5% Ethanol), diamide (2% dextrose, 0.5% ammonium sulphate, 0.85 mM, 30°C), or 37°C (2% dextrose, 0.5% ammonium sulphate, 37°C). We collected samples of ~10^8 cells at eight time points over a period of 36 hours and stored them at -80°C. Cultures were maintained in log phase by regular dilution with fresh media, maintaining a population size of 10^9 or greater throughout the bulk competition. Bulk competition from the standard condition were conducted in technical duplicates from the frozen yeast library.

DNA Preparation and Sequencing

We isolated plasmid DNA from each bulk competition time point as described [32]. Purified plasmid was linearized with AscI. Barcodes were amplified by 19 cycles of PCR using Phusion polymerase (NEB) and primers that add Illumina adapter sequences and an 8 bp identifier sequence used to distinguish libraries and time points. The identifier sequence was located at positions 91-98 relative to the illumine primer and the barcode was located at positions 1-18. PCR products were purified two times over silica columns (Zymo Research) and quantified using the KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) on a Bio-Rad CFX machine. Samples were pooled and sequenced on an Illumina NextSeq instrument in single-end 100 bp mode.

Analysis of Bulk Competition Sequencing Data

Illumina sequence reads were filtered for Phred scores >20 and strict matching of the sequence to the expected template and identifier sequence. Reads that passed these filters were parsed based on the identifier sequence. For each condition/time-point identifier, each unique N18 read was counted. The unique N18 count file was then used to identify the frequency of each mutant using the variant-barcode association table. To generate a cumulative count for each

codon and amino acid variant in the library, the counts of each associated barcode were summed.

Determination of Selection Coefficient

The frequency of each variant in the library relative to wildtype synonyms was determined at each time point and the slope of the log of this ratio versus generation time was used to calculate the raw selection coefficient. This procedure sets the selection coefficient for the average wildtype synonym to neutral (s=0). Selection coefficients (s) were scaled so that the average stop codon in each environmental condition represented a null allele (s=-1). For the second replicate in standard conditions, we noted a small fitness defect (s≈-0.05) for wildtype synonyms at positions 679-709 relative to other positions. We do not understand the source of this behavior, and chose to normalize to wildtype synonyms from 1-678 for this condition and to exclude positions 679-709 from analyses that include the second replicate of standard conditions. We did not observe this behavior in any other condition.

Yeast growth analysis

Individual Hsp90 variants were generated and analyzed essentially as previously described [32]. Variants were generated by site directed mutagenesis and transformed into DBY288 cells. Selected transformed colonies were grown in liquid SRGal-W media to mid-log phase at 30°C, washed three times and grown in shutoff media (SD-W) at either 30C or 37C. After sufficient time to stall the growth of control cells lacking a rescue copy of Hsp90 (~16 hours), cell density was monitored based on absorbance at 600 nm over time and fit to an exponential growth curve to quantify growth rate.

Natural variation in Hsp90 sequence

We analyzed sequence variation in a previously described alignment of Hsp90 protein sequences from 261 eukaryotic species that broadly span a billion years of evolutionary distance [67].

REFERENCES

1. Darwin C: *On the origin of species by means of natural selection.* John Murray, London (1859).

2. Darwin C, Wallace A: **On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection.** *Journal of hte Proceedings of the Linnean Society of London* (1858) **Zoology**(3):45-62.

3. Wright S: **The roles of mutation, inbreeding, crossbreeding, and selection in evolution.** *Proceedings of the Sixth International Congress on Genetics* (1932) 355-366.

4. Ohta T: **Slightly deleterious mutant substitutions in evolution.** *Nature* (1973) **246**(5428):96-98.

5. Tutt JW: *British moths.* Routledge, London (1896).

6. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN: **Genome-wide association studies for complex traits: Consensus, uncertainty and challenges.** *Nature reviews Genetics* (2008) **9**(5):356-369.

7. Jensen JD, Payseur BA, Stephan W, Aquadro CF, Lynch M, Charlesworth D, Charlesworth B: **The importance of the neutral theory in 1968 and 50 years on: A response to kern and hahn 2018.** *Evolution; international journal of organic evolution* (2019) **73**(1):111-114.

8. Kimura M: **Evolutionary rate at the molecular level.** *Nature* (1968) **217**(5129):624-626.

9. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C: **Genomic scans for selective sweeps using snp data.** *Genome research* (2005) **15**(11):1566-1575.

10. Lenski RE, Travisano M: **Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations.** *Proceedings of the National Academy of Sciences of the United States of America* (1994) **91**(15):6808-6814.

11. Barrick JE, Lenski RE: **Genome dynamics during experimental evolution.** *Nature reviews Genetics* (2013) **14**(12):827-839.

12. Allison AC: **Protection afforded by sickle-cell trait against subtertian malareal infection.** *British medical journal* (1954) **1**(4857):290-294.

13. Barrett RDH, Laurent S, Mallarino R, Pfeifer SP, Xu CCY, Foll M, Wakamatsu K, Duke-Cohan JS, Jensen JD, Hoekstra HE: **Linking a mutation to survival in wild mice.** *Science* (2019) **363**(6426):499-504.

14. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP *et al*: **Functional profiling of the saccharomyces cerevisiae genome.** *Nature* (2002) **418**(6896):387-391.

15. Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, Baker D, Fields S: **High-resolution mapping of protein sequence-function relationships.** *Nature methods* (2010) **7**(9):741-746.

16. Boucher JI, Bolon DN, Tawfik DS: **Quantifying and understanding the fitness effects of protein mutations: Laboratory versus nature.** *Protein science : a publication of the Protein Society* (2016) **25**(7):1219-1226.

17. Boucher JI, Cote P, Flynn J, Jiang L, Laban A, Mishra P, Roscoe BP, Bolon DN: **Viewing protein fitness landscapes through a next-gen lens.** *Genetics* (2014) **198**(2):461-471.

18. Hietpas RT, Jensen JD, Bolon DN: **Experimental illumination of a fitness landscape.** *Proc Natl Acad Sci U S A* (2011) **108**(19):7896-7901.

19. Mavor D, Barlow K, Thompson S, Barad BA, Bonny AR, Cario CL, Gaskins G, Liu Z, Deming L, Axen SD, Caceres E *et al*: **Determination of ubiquitin fitness landscapes under different chemical stresses in a classroom setting.** *eLife* (2016) **5**(

20. Stiffler MA, Hekstra DR, Ranganathan R: **Evolvability as a function of purifying selection in tem-1 beta-lactamase.** *Cell* (2015) **160**(5):882-892.

21. Firnberg E, Labonte JW, Gray JJ, Ostermeier M: **A comprehensive, high-resolution map of a gene's fitness landscape.** *Molecular biology and evolution* (2014) **31**(6):1581-1592.

22. Jiang L, Liu P, Bank C, Renzette N, Prachanronarong K, Yilmaz LS, Caffrey DR, Zeldovich KB, Schiffer CA, Kowalik TF, Jensen JD *et al*: **A balance between inhibitor binding and substrate processing confers influenza drug resistance.** *Journal of molecular biology* (2016) **428**(3):538-553.

23. Dingens AS, Arenz D, Overbaugh J, Bloom JD: **Massively parallel profiling of hiv-1 resistance to the fusion inhibitor enfuvirtide.** *Viruses* (2019) **11**(5).

24. Doud MB, Lee JM, Bloom JD: **How single mutations affect viral escape from broad and narrow antibodies to h1 influenza hemagglutinin.** *Nature communications* (2018) **9**(1):1386.

25. Dandage R, Pandey R, Jayaraj G, Rai M, Berger D, Chakraborty K: **Differential strengths of molecular determinants guide environment specific mutational fates.** *PLoS genetics* (2018) **14**(5):e1007419.

26. Lindquist S: **Regulation of protein synthesis during heat shock.** *Nature* (1981) **293**(5830):311-314.

27. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO: **Genomic expression programs in the response of yeast cells to environmental changes.** *Molecular biology of the cell* (2000) **11**(12):4241-4257.

28. Krukenberg KA, Street TO, Lavery LA, Agard DA: **Conformational dynamics of the molecular chaperone hsp90.** *Q Rev Biophys* (2011) **44**(2):229-255.

29. Taipale M, Krykbaeva I, Koeva M, Kayatekin C, Westover KD, Karras GI, Lindquist S: **Quantitative analysis of hsp90-client interactions reveals principles of substrate recognition.** *Cell* (2012) **150**(5):987-1001.

30. Hietpas RT, Bank C, Jensen JD, Bolon DNA: **Shifting fitness landscapes in response to altered environments.** *Evolution; international journal of organic evolution* (2013) **67**(12):3512-3522.

31. Hietpas R, Roscoe B, Jiang L, Bolon DN: **Fitness analyses of all possible point mutations for regions of genes in yeast.** *Nature protocols* (2012) **7**(7):1382-1396.

32. Jiang L, Mishra P, Hietpas RT, Zeldovich KB, Bolon DN: **Latent effects of hsp90 mutants revealed at reduced expression levels.** *PLoS Genet* (2013) **9**(6):e1003600.

33. Bank C, Hietpas RT, Wong A, Bolon DN, Jensen JD: **A bayesian mcmc approach to assess the complete distribution of fitness effects of new mutations: Uncovering the potential for adaptive walks in challenging environments.** *Genetics* (2014) **196**(3):841-852.

34. Fragata I, Blanckaert A, Dias Louro MA, Liberles DA, Bank C: **Evolution in the light of fitness landscape theory.** *Trends in ecology & evolution* (2019) **34**(1):69-82.

35. Mishra P, Flynn JM, Starr TN, Bolon DNA: **Systematic mutant analyses elucidate general and client-specific aspects of hsp90 function.** *Cell reports* (2016) **15**(3):588-598.

36. Rohl A, Rohrberg J, Buchner J: **The chaperone hsp90: Changing partners for demanding clients.** *Trends Biochem Sci* (2013) **38**(5):253-262.

37. Piper PW: **The heat shock and ethanol stress responses of yeast exhibit extensive similarity and functional overlap.** *FEMS microbiology letters* (1995) **134**(2-3):121-127.

38. Picard D, Khursheed B, Garabedian MJ, Fortin MG, Lindquist S, Yamamoto KR: **Reduced levels of hsp90 compromise steroid receptor action in vivo.** *Nature* (1990) **348**(6297):166-168.

39. Kacser H, Burns JA: **The molecular basis of dominance.** *Genetics* (1981) **97**(3-4):639-666.

40. Dykhuizen DE, Dean AM, Hartl DL: **Metabolic flux and fitness.** *Genetics* (1987) **115**(1):25-31.

41. Bershtein S, Mu W, Serohijos AW, Zhou J, Shakhnovich EI: **Protein quality control acts on folding intermediates to shape the effects of mutations on organismal fitness.** *Molecular cell* (2013) **49**(1):133-144.

42. Ali MM, Roe SM, Vaughan CK, Meyer P, Panaretou B, Piper PW, Prodromou C, Pearl LH: **Crystal structure of an hsp90-nucleotide-p23/sba1 closed chaperone complex.** *Nature* (2006) **440**(7087):1013-1017.

43. Zhang M, Kadota Y, Prodromou C, Shirasu K, Pearl LH: **Structural basis for assembly of hsp90-sgt1-chord protein complexes: Implications for chaperoning of nlr innate immunity receptors.** *Mol Cell* (2010) **39**(2):269-281.

44. Verba KA, Wang RY, Arakawa A, Liu Y, Shirouzu M, Yokoyama S, Agard DA: **Atomic structure of hsp90-cdc37-cdk4 reveals that hsp90 traps and stabilizes an unfolded kinase.** *Science* (2016) **352**(6293):1542-1547.

45. Meyer P, Prodromou C, Hu B, Vaughan C, Roe SM, Panaretou B, Piper PW, Pearl LH: **Structural and functional analysis of the middle segment of hsp90: Implications for atp**

hydrolysis and client protein and cochaperone interactions. *Mol Cell* (2003) **11**(3):647-658.

46. Roe SM, Ali MM, Meyer P, Vaughan CK, Panaretou B, Piper PW, Prodromou C, Pearl LH: **The mechanism of hsp90 regulation by the protein kinase-specific cochaperone p50(cdc37).** *Cell* (2004) **116**(1):87-98.

47. Hagn F, Lagleder S, Retzlaff M, Rohrberg J, Demmer O, Richter K, Buchner J, Kessler H: **Structural analysis of the interaction between hsp90 and the tumor suppressor protein p53.** *Nat Struct Mol Biol* (2011) **18**(10):1086-1093.

48. Hawle P, Siepmann M, Harst A, Siderius M, Reusch HP, Obermann WM: **The middle domain of hsp90 acts as a discriminator between different types of client proteins.** *Mol Cell Biol* (2006) **26**(22):8385-8395.

49. Bohen SP, Yamamoto KR: **Isolation of hsp90 mutants by screening for decreased steroid receptor function.** *Proc Natl Acad Sci U S A* (1993) **90**(23):11424-11428.

50. Lorenz OR, Freiburger L, Rutz DA, Krause M, Zierer BK, Alvira S, Cuellar J, Valpuesta JM, Madl T, Sattler M, Buchner J: **Modulation of the hsp90 chaperone cycle by a stringent client protein.** *Mol Cell* (2014) **53**(6):941-953.

51. Kravats AN, Hoskins JR, Reidy M, Johnson JL, Doyle SM, Genest O, Masison DC, Wickner S: **Functional and physical interaction between yeast hsp90 and hsp70.** *Proc Natl Acad Sci U S A* (2018) **115**(10):E2210-E2219.

52. Genest O, Reidy M, Street TO, Hoskins JR, Camberg JL, Agard DA, Masison DC, Wickner S: **Uncovering a region of heat shock protein 90 important for client binding in e. Coli and chaperone function in yeast.** *Mol Cell* (2013) **49**(3):464-473.

53. Retzlaff M, Stahl M, Eberl HC, Lagleder S, Beck J, Kessler H, Buchner J: **Hsp90 is regulated by a switch point in the c-terminal domain.** *EMBO reports* (2009) **10**(10):1147-1153.

54. Nathan DF, Lindquist S: **Mutational analysis of hsp90 function: Interactions with a steroid receptor and a protein kinase.** *Mol Cell Biol* (1995) **15**(7):3917-3925.

55. Meyer P, Prodromou C, Liao C, Hu B, Roe SM, Vaughan CK, Vlasic I, Panaretou B, Piper PW, Pearl LH: **Structural basis for recruitment of the atpase activator aha1 to the hsp90 chaperone machinery.** *EMBO J* (2004) **23**(6):1402-1410.

56. Wayne N, Bolon DN: **Dimerization of hsp90 is required for in vivo function. Design and analysis of monomers and dimers.** *J Biol Chem* (2007) **282**(48):35386-35395.

57. Lynch M: **Evolutionary diversification of the multimeric states of proteins.** *Proceedings of the National Academy of Sciences of the United States of America* (2013) **110**(30):E2821-2828.

58. Chakravarty S, Varadarajan R: **Residue depth: A novel parameter for the analysis of protein structure and stability.** *Structure* (1999) **7**(7):723-732.

59. Borkovich KA, Farrelly FW, Finkelstein DB, Taulien J, Lindquist S: **Hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures.** *Mol Cell Biol* (1989) **9**(9):3919-3930.

60. Fisher RA: **The genetical theory of natural selection.** *Clarendon Press, Oxford* (1930).

61. Martin G, Lenormand T: **A general multivariate extension of fisher's geometrical model and the distribution of mutation fitness effects across species.** *Evolution; international journal of organic evolution* (2006) **60**(5):893-907.

62. Hietpas RT, Bank, C., Jensen, J.D and Bolon, D. N: **Shifting fitness landscapes in response to altered enviornments.** In: *Evolution.* (2013).

63. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y: **An evolutionarily conserved mechanism for controlling the efficiency of protein translation.** *Cell* (2010) **141**(2):344-354.

64. Li GW: **How do bacteria tune translation efficiency?** *Current opinion in microbiology* (2015) **24**(66-71.

65. Plotkin JB, Kudla G: **Synonymous but not the same: The causes and consequences of codon bias.** *Nature reviews Genetics* (2011) **12**(1):32-42.

66. Lang GI, Desai MM: **The spectrum of adaptive mutations in experimental evolution.** *Genomics* (2014) **104**(6 Pt A):412-416.

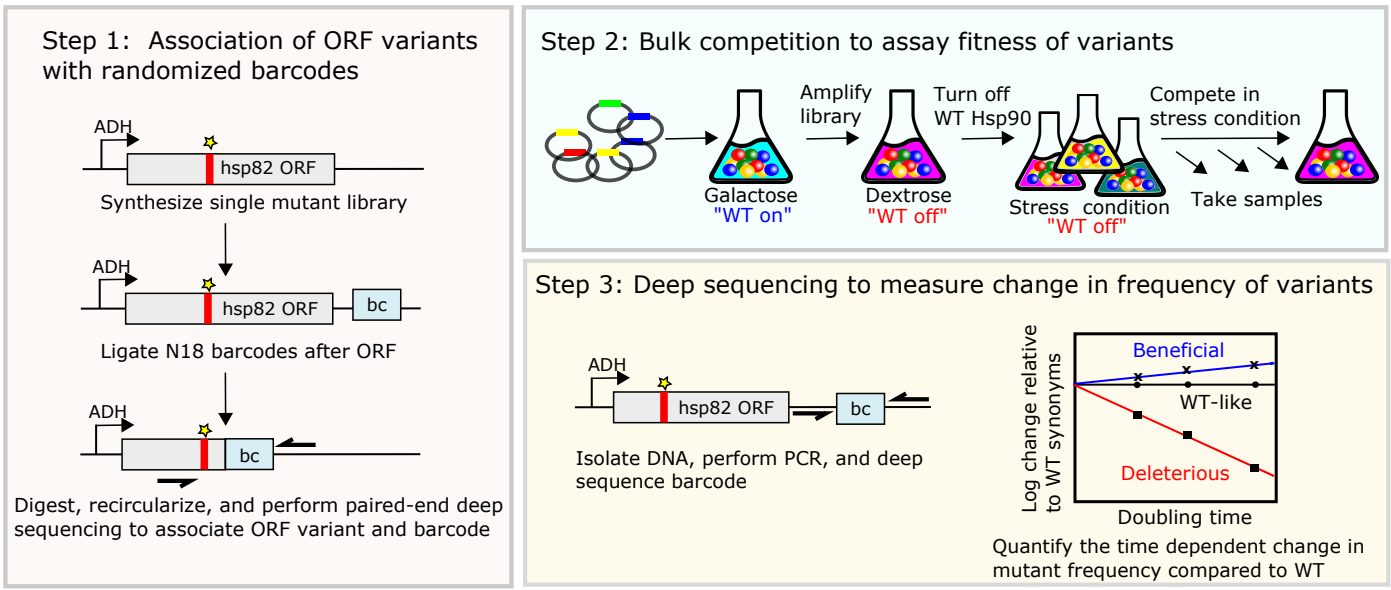
67. Starr TN, Flynn JM, Mishra P, Bolon DNA, Thornton JW: **Pervasive contingency and entrenchment in a billion years of hsp90 evolution.** *Proceedings of the National Academy of Sciences of the United States of America* (2018) **115**(17):4453-4458.

68. Mustonen V, Lassig M: **From fitness landscapes to seascapes: Non-equilibrium dynamics of selection and adaptation.** *Trends in genetics : TIG* (2009) **25**(3):111-119.

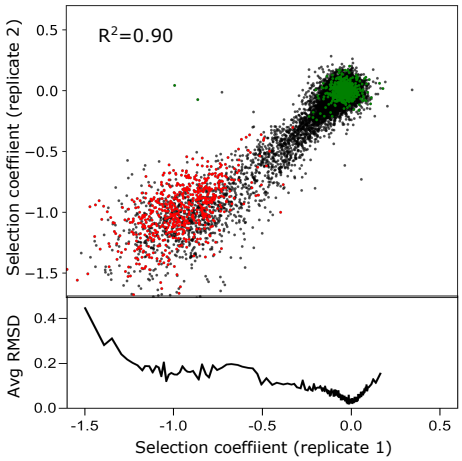
69. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R: **Interactivenn: A web-based tool for the analysis of sets through venn diagrams.** *BMC bioinformatics* (2015) **16**(169.

Figure 1

A.


Step 1: Association of ORF variants with randomized barcodes
ADH
hsp82 ORF
Synthesize single mutant library
ADH
hsp82 ORF
bc
Ligate N18 barcodes after ORF
ADH
bc
Digest, recircularize, and perform paired-end deep sequencing to associate ORF variant and barcode
Step 2: Bulk competition to assay fitness of variants
Amplify library
Turn off WT Hsp90
Compete in stress condition
Galactose
"WT on"
Dextrose
"WT off"
Stress condition
"WT off"
Take samples
Step 3: Deep sequencing to measure change in frequency of variants
ADH
hsp82 ORF
bc
Isolate DNA, perform PCR, and deep sequence barcode
Log change relative to WT synonyms
Beneficial
WT-like
Deleterious
Doubling time
Quantify the time dependent change in mutant frequency compared to WT

B.


R^2=0.90
Selection coeffiient (replicate 2)
0.5
0.0
−0.5
−1.0
−1.5
Avg RMSD
0.4
0.2
0.0
-1.5
-1.0
-0.5
0.0
0.5
Selection coeffiient (replicate 1)

C.

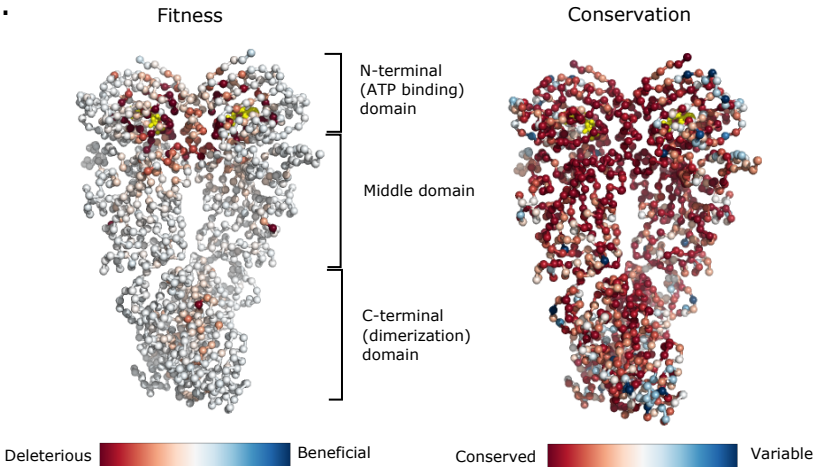
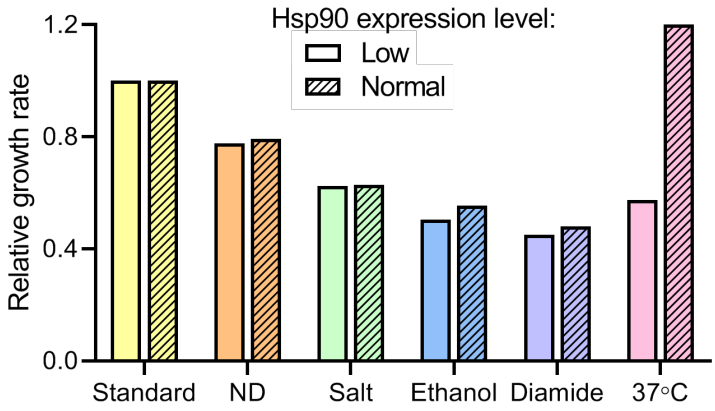
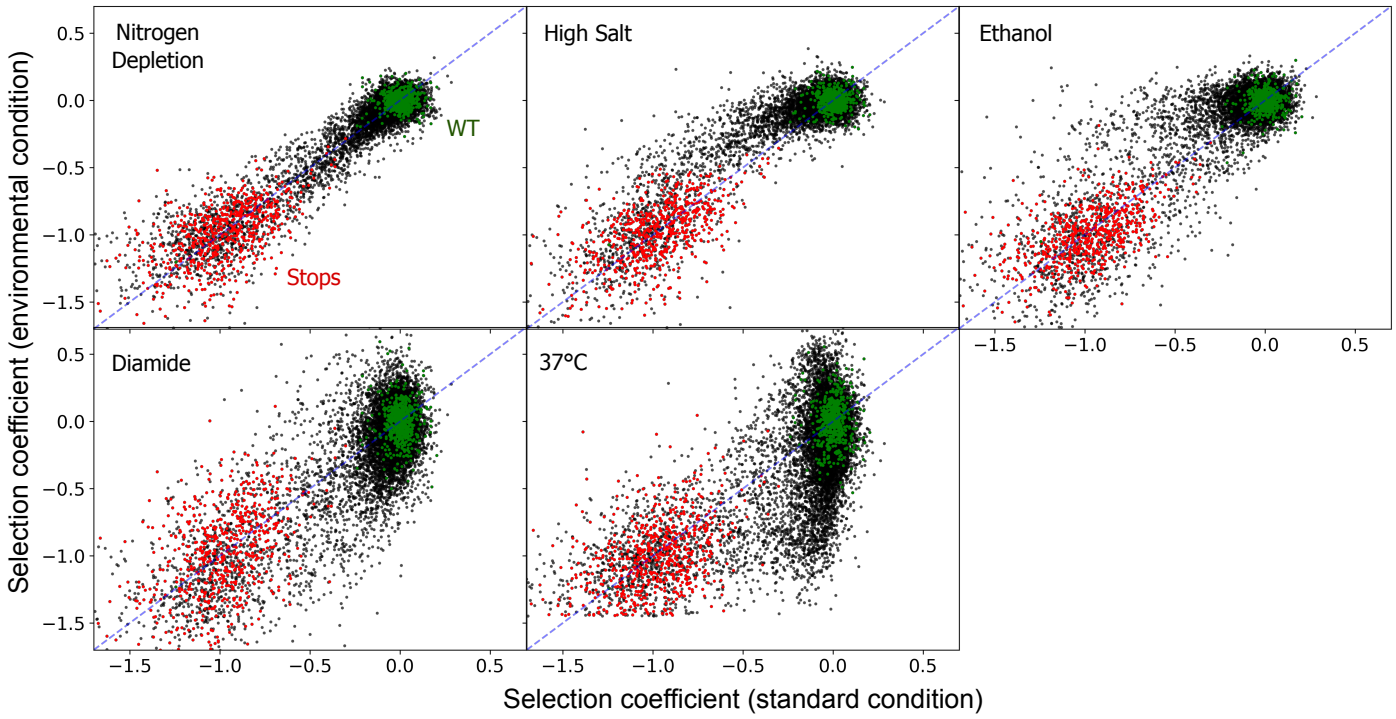

Fitness
Conservation
N-terminal (ATP binding) domain
Middle domain
C-terminal (dimerization) domain
Deleterious
Beneficial
Conserved
Variable

Figure 2

A.


Hsp90 expression level:
Low
Normal
Relative growth rate
1.2
0.8
0.4
0.0
Standard
ND
Salt
Ethanol
Diamide
37°C

B.


Nitrogen Depletion
High Salt
Ethanol
Diamide
37°C
WT
Stops
Selection coefficient (environmental condition)
Selection coefficient (standard condition)
0.5
0.0
−0.5
−1.0
−1.5

C.

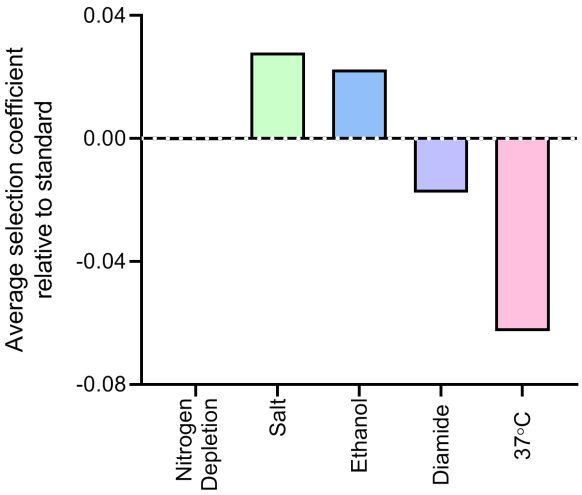
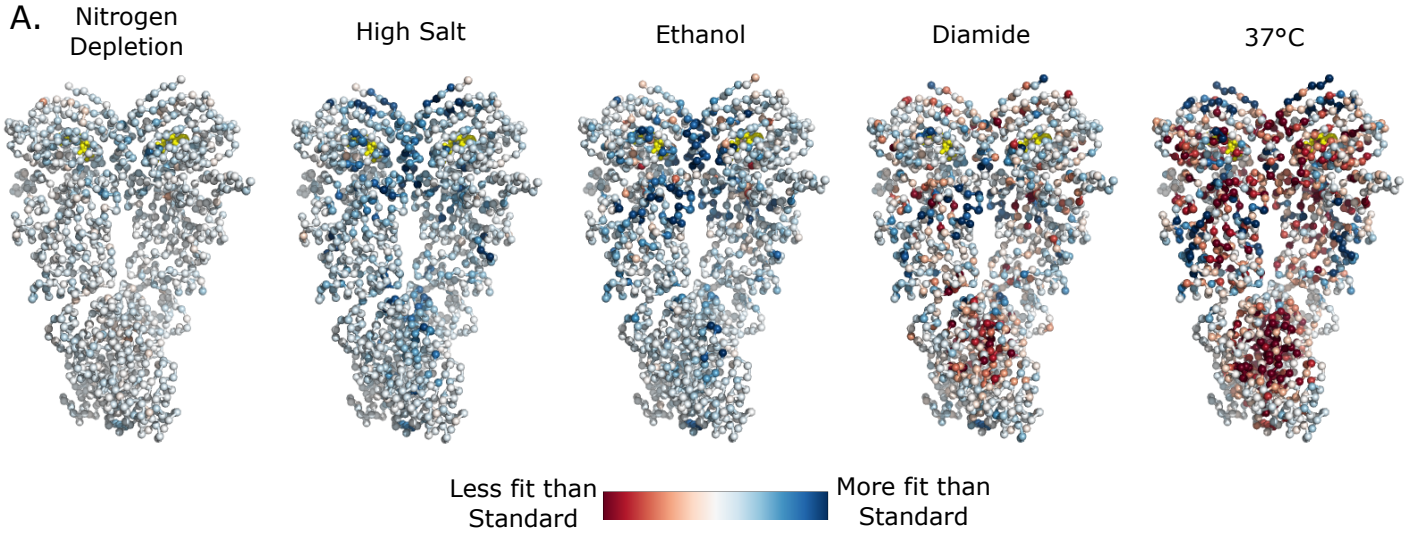

Average selection coefficient relative to standard
0.04
0.00
-0.04
-0.08
Nitrogen Depletion
Salt
Ethanol
Diamide
37°C

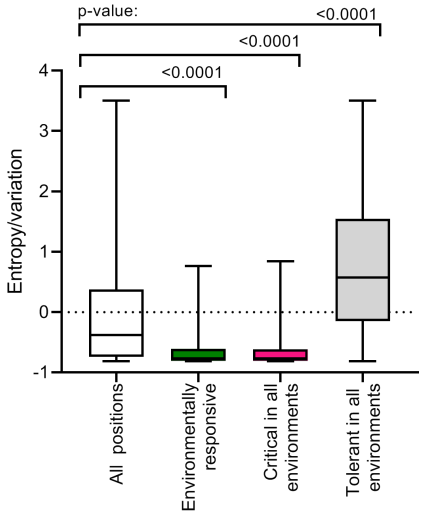
Figure 3

A.

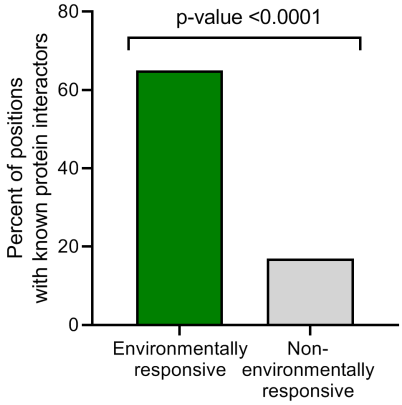


B.

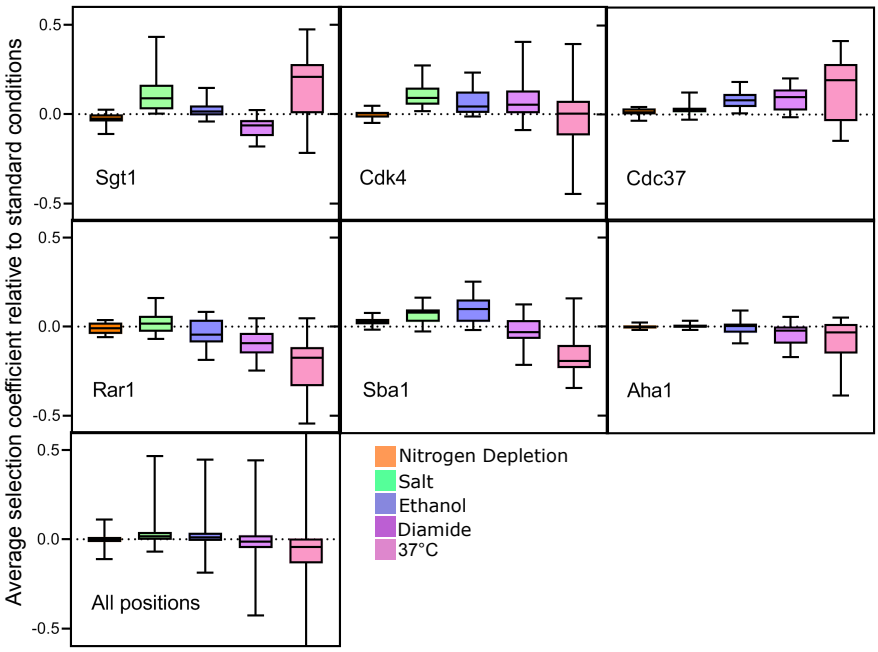
C.



D.



E.



F.

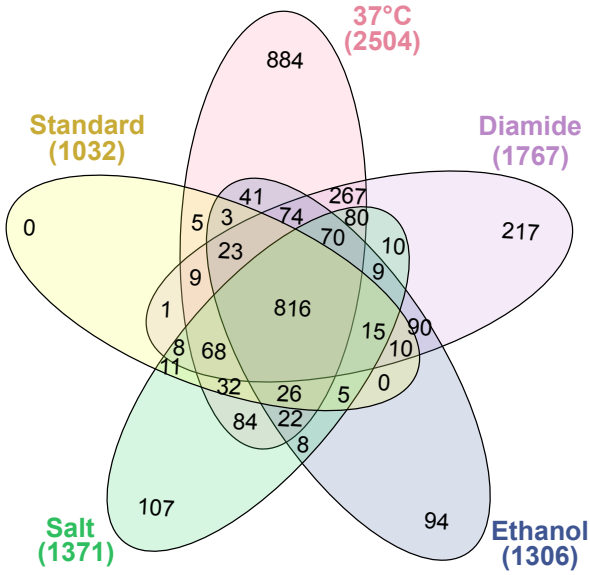
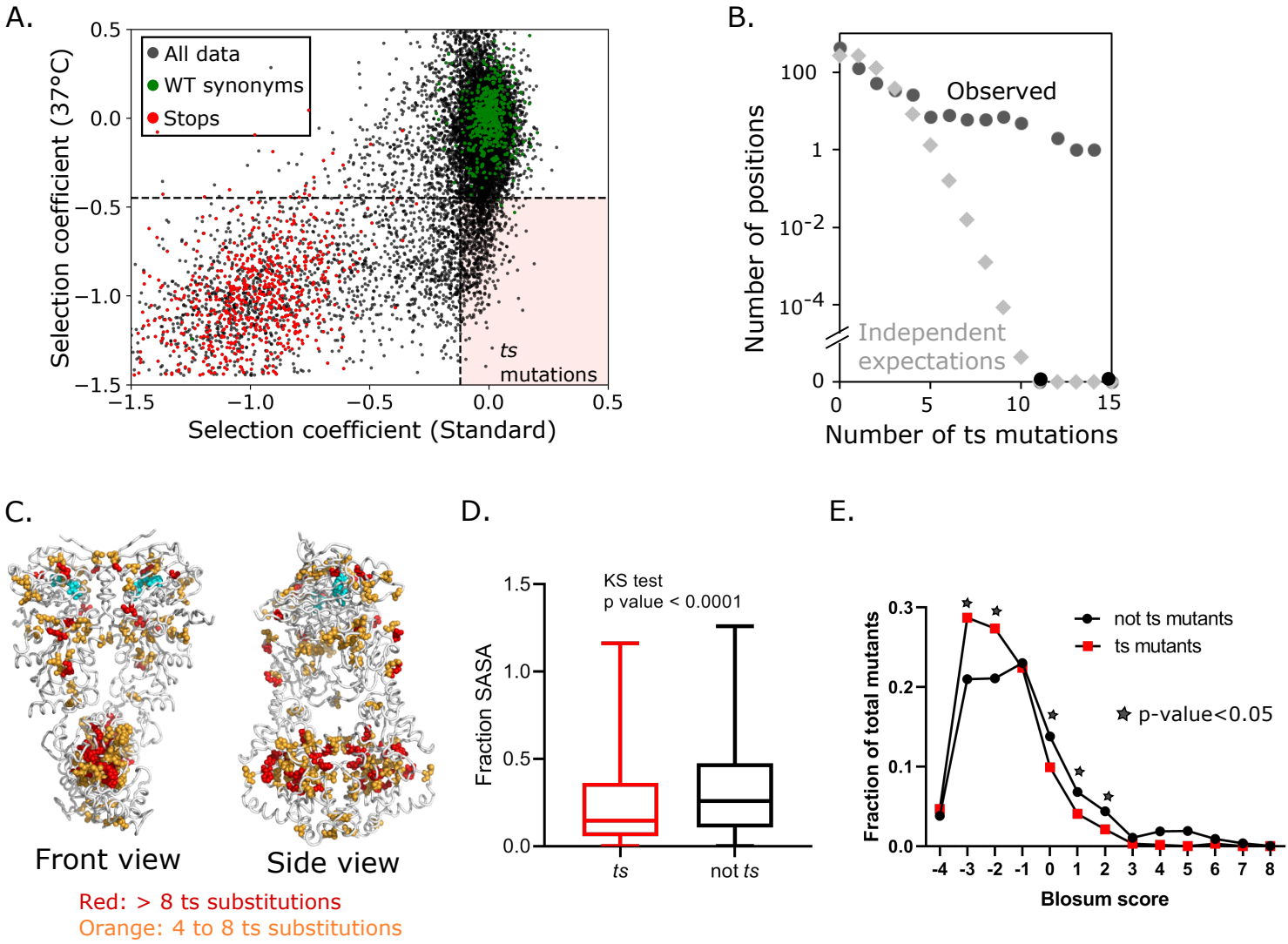


Figure 4


A.
All data
WT synonyms
Stops
Selection coefficient (37°C)
Selection coefficient (Standard)
ts mutations
B.
Observed
Independent expectations
Number of positions
Number of ts mutations
C.
Front view
Side view
Red: > 8 ts substitutions
Orange: 4 to 8 ts substitutions
D.
KS test
p value < 0.0001
Fraction SASA
ts
not ts
E.
not ts mutants
ts mutants
p-value<0.05
Fraction of total mutants
Blosum score

F.

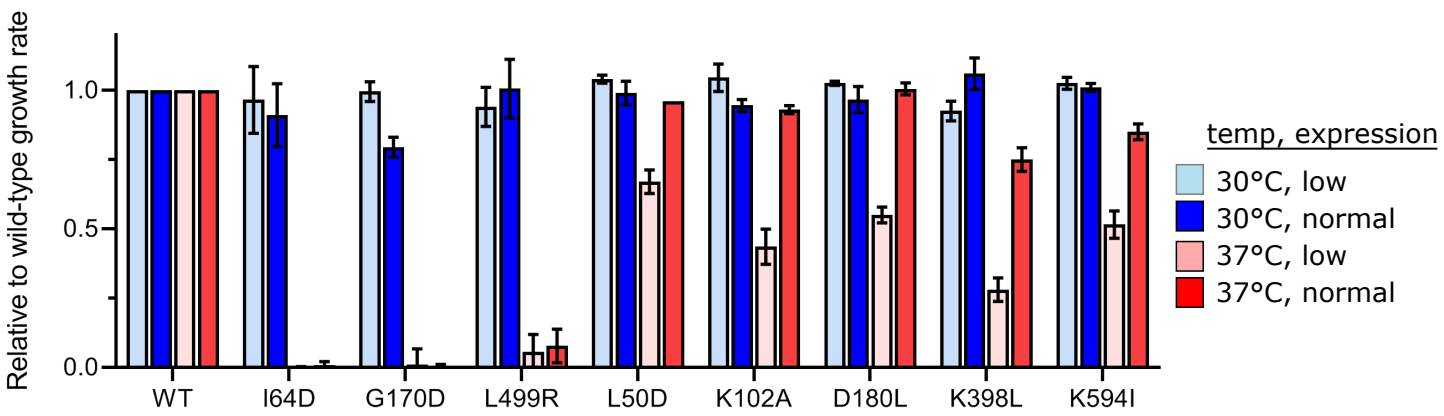

Relative to wild-type growth rate
WT
I64D
G170D
L499R
L50D
K102A
D180L
K398L
K594I
temp, expression
30°C, low
30°C, normal
37°C, low
37°C, normal

Figure 5

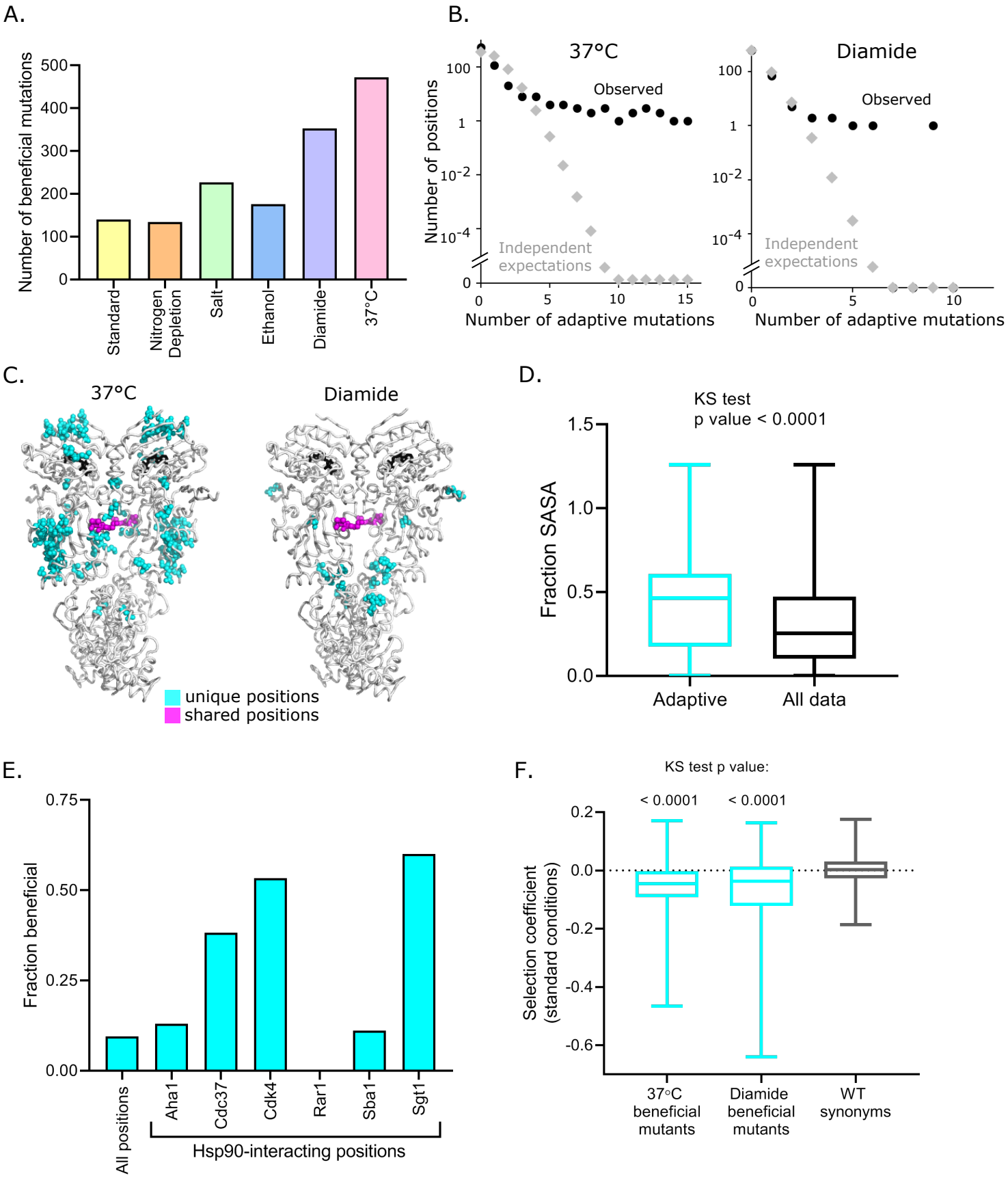
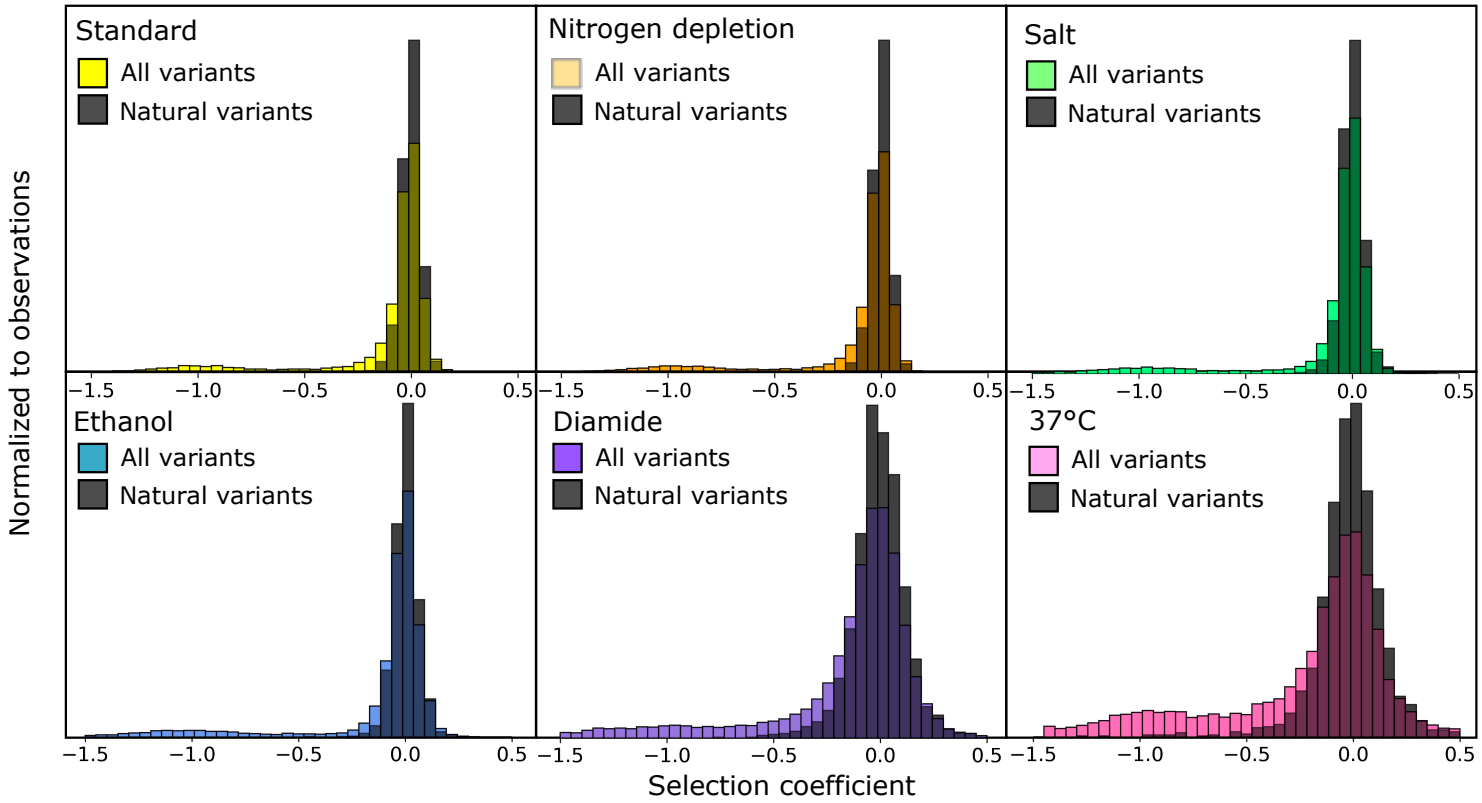
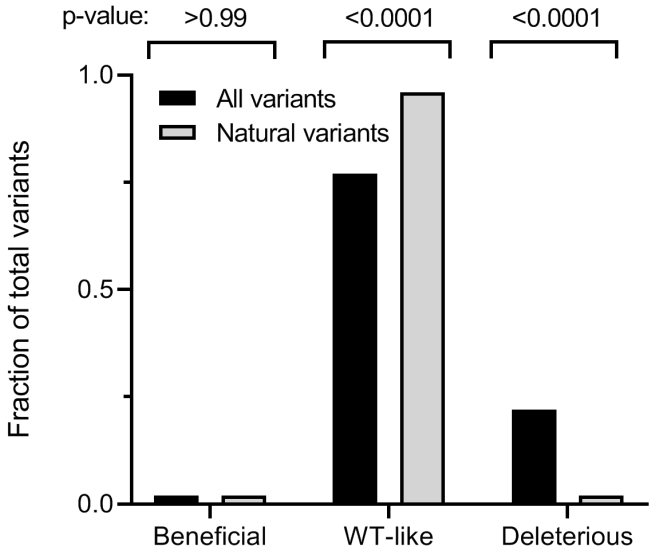

A.
Number of beneficial mutations
500
400
300
200
100
0
Standard
Nitrogen Depletion
Salt
Ethanol
Diamide
37°C
B.
37°C
Diamide
Observed
Independent expectations
Number of positions
100
1
10^{-2}
10^{-4}
0
0
5
10
15
Number of adaptive mutations
C.
37°C
Diamide
unique positions
shared positions
D.
KS test
p value < 0.0001
Fraction SASA
1.5
1.0
0.5
0.0
Adaptive
All data
E.
Fraction beneficial
0.75
0.50
0.25
0.00
All positions
Aha1
Cdc37
Cdk4
Rar1
Sba1
Sgt1
Hsp90-interacting positions
F.
KS test p value:
< 0.0001
< 0.0001
Selection coefficient (standard conditions)
0.2
0.0
-0.2
-0.4
-0.6
37°C beneficial mutants
Diamide beneficial mutants
WT synonyms

Figure 6

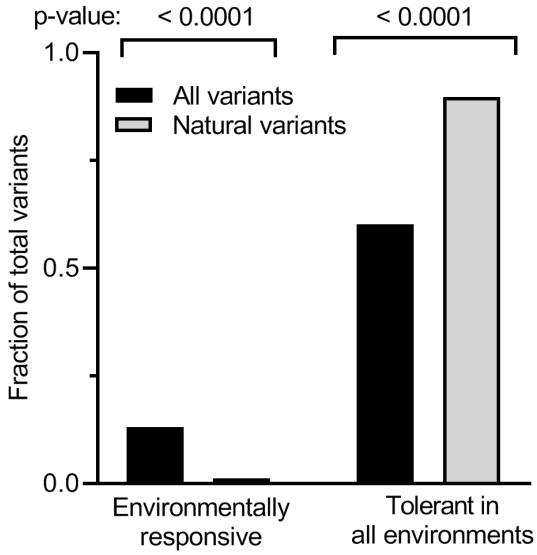
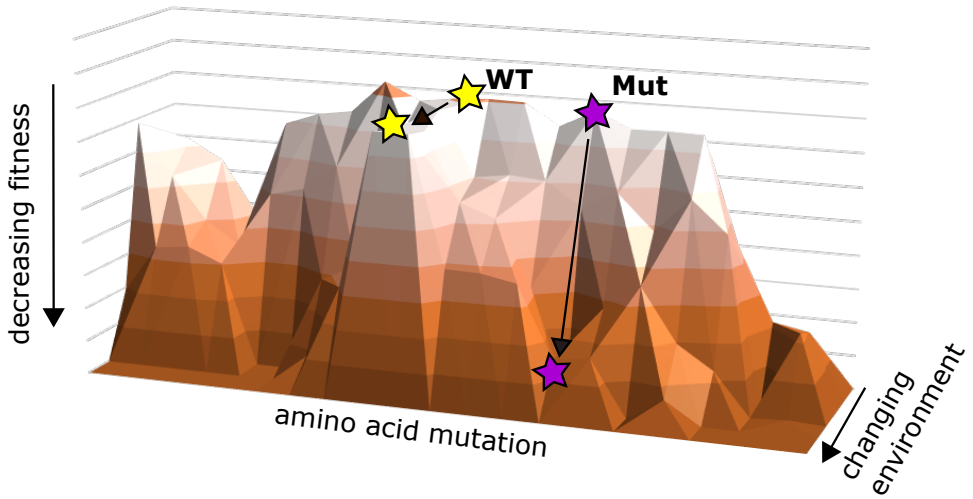
A.



B.



C.



D.