The adaptive potential of the middle domain of yeast Hsp90

Pamela A. Cote-Hammarlof[1]*, Inês Fragata[2*§], Julia Flynn[1], David Mavor[1] , Konstantin B. Zeldovich[1], Claudia Bank[2#], and Daniel N.A. Bolon[1#]

*Co-first authors

#Co-corresponding authors

[1] University of Massachusetts Medical School, Worcester, USA

[2] Instituto Gulbenkian de Ciência, Oeiras, Portugal

[§] Current address: Centre for Ecology and Evolution, Lisbon, Portugal

Corresponding authors:

Daniel N.A. Bolon, Dan.Bolon@umassmed.edu

Claudia Bank, evoldynamics@gmail.com

Abstract (up to 250 words, currently 249)

Comparing the distribution of fitness effects (DFE) of new mutations across different environments quantifies the potential for adaptation in a given environment and its cost in others. So far, results regarding the cost of adaptation across environments have been mixed, and there were no sufficiently large data sets to map its variation along the genome. Here, we study the DFEs of ≈2300 amino-acid changing mutations obtained from deep mutational scanning of 119 amino acids in the middle domain of heat-shock protein Hsp90 in five environments and at two expression levels. This region is known to be important for client binding, stabilization of the Hsp90 dimer, stabilization of the N terminal-Middle and Middle-C terminal interdomains, and regulation of ATPase-chaperone activity. Interestingly, we find that fitness correlates well across diverse and stressful environments, with the exception of one environment, diamide. Consistent with these results, we find very little cost of adaptation; on average only one in seven beneficial mutations is deleterious in another environment. We identify a hotspot of beneficial mutations in a region of the protein that is located within an allosteric center. The identified protein regions that are enriched in beneficial, deleterious, and costly mutations coincide with residues that are involved in the stabilization of Hsp90 interdomains and stabilization of client binding interfaces, or residues that are involved in ATPase chaperone activity of Hsp90. Thus, our study yields information

regarding the role and adaptive potential of a protein sequence that complements and extends known structural information.

Introduction

The distribution of fitness effects (DFE) describes the proportions of new mutations that are beneficial, deleterious or neutral (Eyre-Walker & Keightley 2007, Loewe 2010, Bataillon & Bailey 2014). It provides a snapshot of the robustness of the genome to changes in the DNA and carries information about the expected amount of genetic diversity within populations. Moreover, the beneficial part of the DFE informs on the adaptive potential of populations when introduced in a new environment (Sniegowski & Gerrish 2010, Bataillon & Bailey 2014). While maintaining the same shape of the DFE, beneficial mutations in one environment can be deleterious in another (Bataillon et al 2011). So far, it has been difficult to address the prevalence of such costs of adaptation (also termed antagonistic pleiotropy), because it was difficult to measure the fitness of the same mutations across various environments (but see, e.g., Hietpas et al 2013, Sane et al 2018).

Around a decade ago, inference of the complete DFE of a focal genomic region in several model species started to become feasible, using a combination of site-directed mutagenesis and deep sequencing (Fowler et al. 2010, Hietpas et al 2011, 2012, Boucher et al. 2014). These studies usually reported a bimodal DFE with two peaks that represent neutral (or wild-type like) mutations and strongly deleterious mutations, respectively (Acevedo et al 2014, Bank et al 2014, Boucher et al 2014). With improvements of experimental and sequencing technologies, studies of hundreds of thousands of mutant effects have become feasible (Melamed et al 2013, Boucher et al 2014, Doud et al 2016, Sarkisyan et al 2016). However, few studies have quantified the impact of environmental challenges on the shape of these DFEs (but see Hietpas et al 2013). That is partly because inferring the adaptive potential across environments and the cost of beneficial mutations in alternative environments requires a high experimental resolution of selection coefficient estimates.

Experimentally quantifying the DFE across environments is important because natural environments can change due to the alterations of abiotic factors including temperature, salinity, and pH, which can have diverse evolutionary consequences that are directly related to the shape of the DFE (Dhar et al. 2011, Arribas et al. 2014, Mumby and van Woesik 2014,

Brennan et al. 2017). Because environments are in general likely to change faster than genetic adaptation can proceed, the long-term fitness of individuals depends on their ability to deal with a wide range of environmental stresses.

Environmental challenges manifest at many biological levels, including the protein level. For example, increased temperature can cause protein unfolding and aggregation, which can disrupt function and ultimately affect organismal fitness and survival (Richter et al. 2010). Chaperones, such as the heat shock protein Hsp90, help the cellular machinery survive stress conditions (Chen et al. 2018). They are an important line of defense for new and recurring environmental challenges and can also function to increase standing genetic variation that can facilitate adaptation to new stress conditions (Barrett and Schluter 2008, Jarosz and Lindquist 2010, Fitzgerald and Rosenberg 2019). Therefore, it is important to understand how chaperones themselves evolve, and how the selection pressure on the chaperone changes upon exposure to different environmental challenges.

Hsp90 is a chaperone that plays an essential role in protecting cells from environmental stress and that is required at elevated levels for yeast growth at high temperature (Borkovich et al. 1989). Under optimal conditions Hsp90 can suppress the phenotypic consequences of mutations in signal transduction pathways by correctly folding mutant clients (Sangster et al. 2004, Jarosz and Lindquist 2010, Zabinsky et al. 2019). However, in stressful conditions there is an increase in the concentration of stress-induced unfolded and damaged proteins, which compete for Hsp90 chaperone function (Sangster et al. 2004, Taipale et al. 2010). Thus, mutations in protein clients and in Hsp90 itself may have environment-dependent phenotypic and ultimately fitness consequences.

Recent studies using systematic mutagenesis have begun to address how mutations to a strongly conserved client binding site of Hsp90 can impact evolutionary adaptation in yeast (Hietpas et al. 2013). Multiple mutations in a nine amino acid client-binding site of yeast Hsp90 provided a growth advantage under elevated salinity conditions (Hietpas et al. 2013, Bank et al 2014). A recent larger-scale study found that at low Hsp90 expression levels, changes in environment greatly changed the shape of the DFE, and that some environments showed a higher prevalence of both beneficial and deleterious mutations (Flynn et al 2019). However, another previous study proposed that at low expression the fitness effects of mutations, especially deleterious ones, should be larger (Jiang et al 2013). Thus, it is

unknown how much of the observed effect was due to expression level and how much was due to environmental changes.

Here we examined a 119 amino acid region (encompassing positions 291-409, Fig. 1) of the middle domain of yeast Hsp90 (aka Hsp82) at normal expression levels. Several studies have demonstrated that the middle domain of Hsp90 plays a prominent role in client binding (Nathan and Lindquist 1995, Nathan et al. 1997, Meyer et al. 2003, Hawle et al. 2006, Hagn et al. 2011), and suggested that mutations in this region may impact the relative affinity or priority of different clients and physiological pathways with the potential to provide an adaptive benefit (Hagn et al. 2011, Zhang et al. 2005, Sato et al. 2000, Karagoz et al. 2014, Verba et al. 2016, Czemeres et al. 2017).

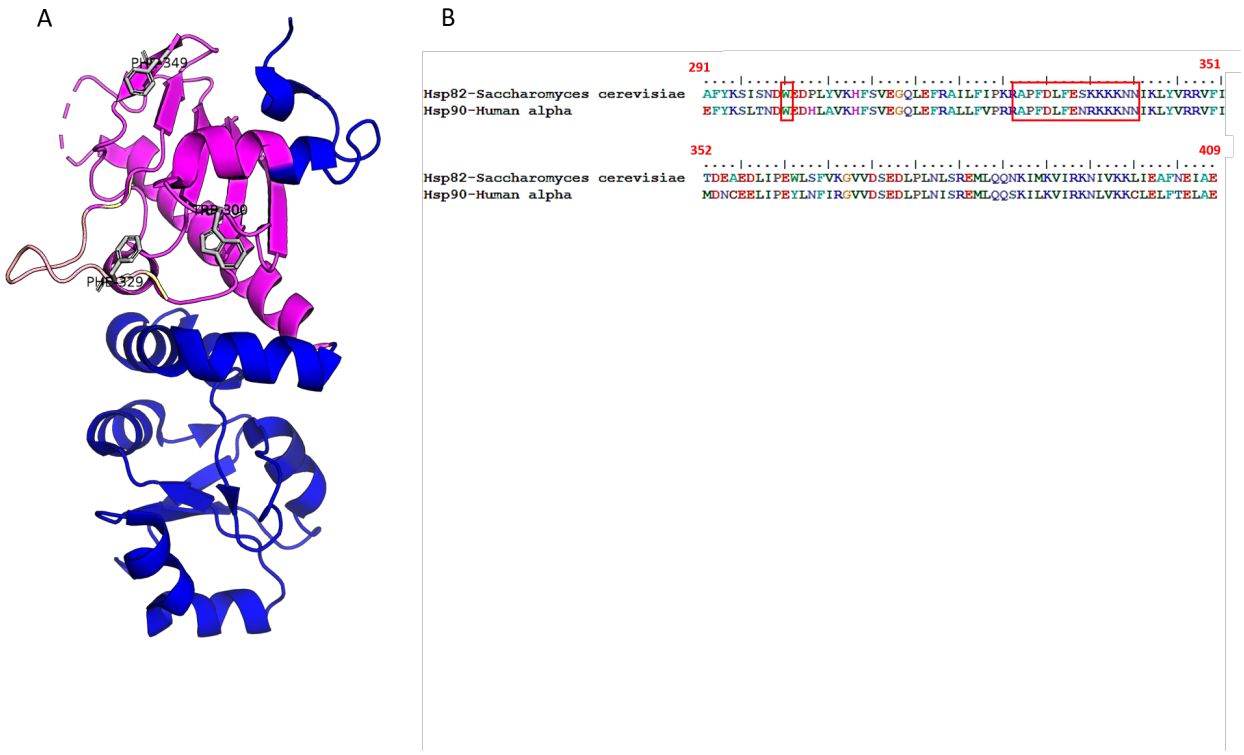


Figure 1: **Middle domain of Hsp90**

(A) Structural representation based on 1HK7.PDB of the middle domain of Hsp90, with amino acids 291-409 that are the focus of this study highlighted in purple. A solvent exposed amphipathic loop implicated in client binding (amino acids 327-341) is highlighted in yellow and residues implicated in client binding (amino acids W300, F329, F349) are shown as gray sticks. (B) Alignment of amino acids 291-409 of yeast Hsp90 with human Hsp90. Boxed in

117 red are amino acid residues implicated in client binding. Amino acids are colored based on
118 their chemical properties.
119
120 We used the EMPIRIC approach of deep mutational scanning (Hietpas et al 2011) to estimate
121 the selection coefficient of all amino-acid changing mutations in the middle domain of yeast
122 Hsp90 across five environments. Whereas the present study was performed at normal
123 expression levels, we included low-expression data from Flynn et al. 2019 in diamide, salt
124 and standard conditions in our analysis to evaluate the impact of protein expression on the
125 fitness effect of mutations. Using the inferred DFEs, we identified regions of the middle
126 domain that stand out with respect to their potential for adaptation upon environmental
127 change. To this end we: a) quantified the impact of environmental changes on the overall
128 shape of the DFE, b) identified regions that show the largest proportions of beneficial or
129 deleterious mutations, respectively, c) quantified hotspots of costs of adaptation, and d)
130 compared the identified regions with known client binding or other structurally important
131 sites to connect the phenotypic and fitness effects of mutations.
132

133 # Results and Discussion

134 **Deep mutational scanning across environments and expression levels reveals**
135 **distributions of fitness effects.** To investigate the adaptive potential of the middle domain of
136 yeast Hsp90 we used systematic site-directed mutagenesis of amino acid positions 291-409
137 that include known client binding sites. This resulted in ≈2300 amino-acid changing
138 mutations, whose selection coefficients were estimated from 2-3 replicates of bulk
139 competitions that were performed in five environments. We focused our analysis on standard
140 lab conditions and four environmental stresses that affect both growth rate (Supplementary
141 figure 1) and the adaptive gene expression response in yeast (Gasch et al. 2000). Two
142 environments caused osmotic stress (0.5M salinity or 0.6M sorbitol) and two caused
143 oxidative stress (0.6 mM H202 or 0.85mM diamide). These conditions were chosen because
144 of the role of Hsp90 in yeast response to these environmental stressors. In yeast, responses to
145 environmental stress require changes in signaling pathways that dramatically alter gene
146 expression programs (de Nadal et al. 2011). For example, during osmotic stress, Hog1, the
147 yeast homolog of the mammalian p38 mitogen activated protein kinase, becomes activated
148 (Hawle et al., 2007) and results in the induction of ~600 genes (Westfall et al. 2004). Hsp90
149 function is required for adaptation to osmotic stress (Yang et al. 2006) and Hsp90 is
150 necessary for the activation of the HOG pathway (Hawle et al. 2007, Yang et al. 2007).
151 Although Hsp90 function is required for the osmotic stress response, increased expression of
152 Hsp90 does not occur in this condition (Gasch et al. 2000). In contrast, the oxidative stresses
153 of diamide and H_20_2 cause an increase in the expression of chaperones, including Hsp90
154 (Gasch et al. 2000). Moreover, while each of these conditions causes changes in the
155 expression of ~900 genes, diamide and H_20_2 stress promote similar gene expression
156 responses to those observed during thermal stress in yeast (Gasch et al. 2000). To allow for
157 comparisons of mutational fitness across expression levels (Dekel and Alon 2005, Manceau
158 et al. 2011), we included three data sets that were obtained in previous work (Flynn et al.
159 2019), which evaluated the same mutations at constitutively low expression levels in the
160 standard lab environment, the high-salinity environment, and the diamide environment.

161 **DFEs across all environments show many wild-type like mutations.** The shape of the
162 DFE indicates the relative importance of purifying or directional selection as compared with
163 neutral evolution. In apparent contrast to the strong conservation of the Hsp90 middle domain
164 in natural populations, we observed DFEs with mostly wild-type like mutations across all
165 environments (Supplementary Figure 2, 3). Here, we categorized mutants as wild-type like if
166 the 95% confidence interval of the estimated selection coefficient overlapped with 0 (see also
167 Supplementary Table 1, Materials & Methods). According to this criterion, at high expression
168 levels, between 50% (in the H_20_2 environment) and 65% (in the standard environment) of
169 mutations showed a fitness effect that is indistinguishable from the reference type. At low
170 expression levels, where the experimental resolution was lower, between 63% and 73% of
171 mutations (in the standard and diamide environments, respectively) showed a wild-type like
172 effect. Large numbers of wild-type like mutations have been observed previously in deep
173 mutational scanning studies of DFEs (Soskine and Tawfik 2010, Hietpas et al 2013, Melamed
174 et al 2013, Bank et al 2014, Hom et al 2019). Both biological and technical factors can be
175 invoked as an explanation for the large number of wild-type like mutations. Firstly, the
176 resolution of the experiment is likely much lower than the resolution at which natural
177 selection may occur in large yeast populations (Ohta 1992, Boucher et al. 2016). Secondly,
178 selection pressures in the laboratory might differ greatly from those in nature (e.g., Reznick
179 & Ghalambor 2005, Kvitek and Sherlock 2013). Finally and relatedly, natural environments
180 might be fundamentally different and rapidly fluctuating (Mustonen and Lässig 2009). For
181 example, a recent study of the DFE of the full Hsp90 sequence found that mutations that were
182 tolerated across a set of diverse environments were those most likely observed in natural
183 sequences (Flynn et al., 2019).

184

185 **Correlations across conditions reveal that diamide stands out with respect to mutational**
186 **effects.** We next quantified the correlations of fitness effects across replicates, environments
187 and expression levels (Figure 2, Supplementary Table 2). Consistent with the high accuracy
188 of the experiment, we observed strong correlations between replicates (mean Pearson
189 correlation r=0.91). Across environments and expression levels, we observed a large
190 variation of correlations ranging from r=0.36 between diamide and low-expression salt to
191 r=0.98 between H_20_2 and the standard environment. In this analysis, diamide clearly stood
192 out by showing consistently lower correlations with other environments and expression levels
193 than all others (mean correlation including diamide r=0.64, mean correlation of all others
194 r=0.83). This could be because diamide exerts multiple negative effects on the yeast cells. For
195 example, among the environments we investigated, diamide was the only condition that
196 induces the expression of cell wall biosynthesis genes and genes involved in protein secretion
197 and processing in the Endoplasmic Reticulum, which indicates its role in cell wall damage
198 (Gasch et al. 2000). Diamide was also shown to affect individual transcription factors
199 differently from H_20_2, the other oxidizing environment in our selection. For example, the
200 Yap1 transcription factor is nuclear localized and active in diamide, but cytoplasmic and
201 inactive in H_20_2 (Gulshan et al. 2011). Further analyses will be required to determine the
202 molecular features of diamide stress that elicit distinct selection on Hsp90 sequence.

203

204

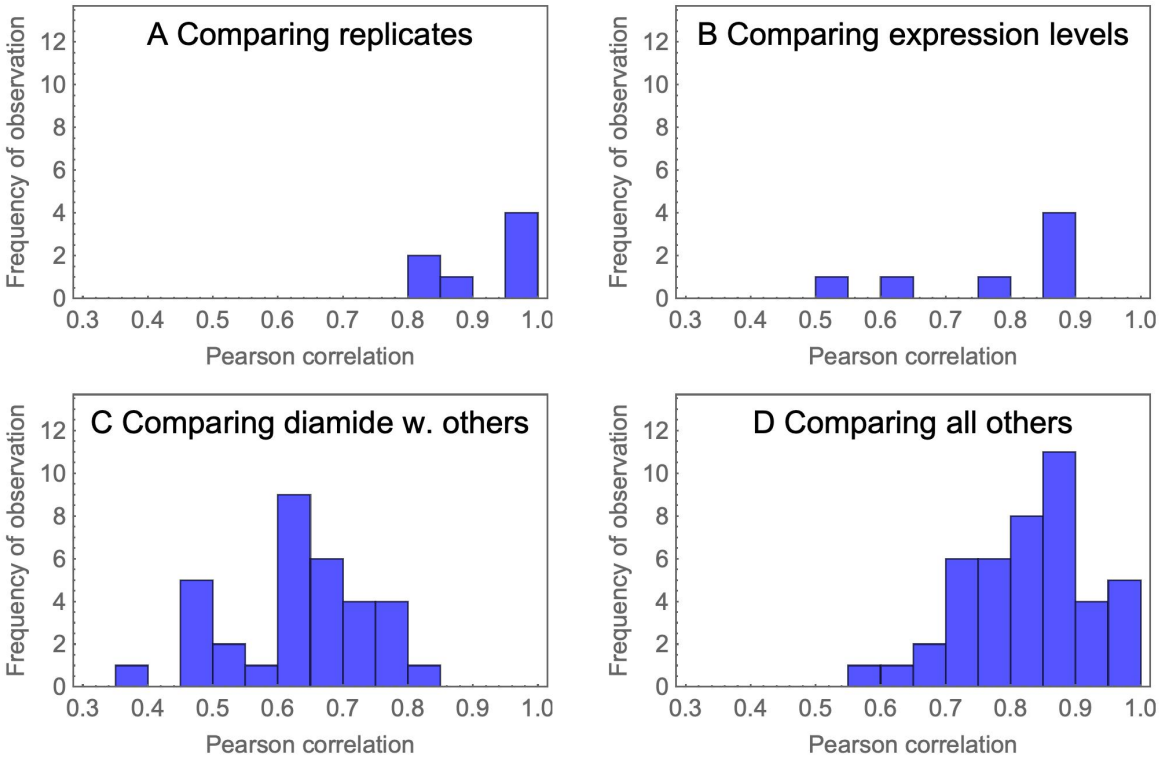


Figure 2: Histogram of correlations of fitness effects among replicates, expression levels, and environments reveals that diamide stands out as different. The histograms correspond to comparisons of replicates (A), expression levels (B), comparisons that include diamide (C), and all other comparisons (D). Correlations between replicates and between several environment pairs are in general high, whereas correlations between diamide and the other environments tend to be lower. Correlations between the same environment at different expression levels vary, with diamide showing the lowest correlation between expression levels (see also Supplementary Table 2). Altogether, this suggests that cells react in unique ways to the oxidative stress enacted by exposure to diamide, and that Hsp90 plays a large role in the response to this stress.

We also observed a tendency for larger correlations of selection coefficients from the same batch of experiments (Supplementary Table 2). Therefore, we performed some of the analyses (as indicated below) using the average of the estimated fitness effects between replicates and re-categorizing mutations based on the mutant category of both replicates (see also Materials and Methods).

Finally, we compared the correlations of fitness effects across environments with those from the same environment, but obtained at a different (constitutive) expression levels. A general expectation is that the Pearson correlation of DFEs between different expression levels should on average be larger than between environmental conditions. This would be consistent with the results and the model of Jiang et al. 2013, who showed that the variance of the DFE increases at lower expression level. That is expected because at lower expression, protein function and fitness become more directly related such that small functional effects should become more visible to growth analyses. However, we see only small differences in the mean correlations within environments at the same expression level compared with correlations between expression levels (mean Pearson correlation across expression levels: r=0.76, vs. r=0.78 for correlations of environments with the same expression level). Again, we find that diamide stands out with a particularly low correlation between low and normal expression levels. This makes sense with respect to the above-mentioned arguments, because an increase in Hsp90 expression levels is required to maintain protein and cellular homeostasis under diamide conditions (Gasch et al. 2000), which is compromised when Hsp90 is expressed at lower levels.

The large correlations observed between most environments suggest that potential costs of adaptation (discussed in detail further below) should be the exception rather than the rule. This is at odds with previous results from a smaller region of Hsp90 (Hietpas et al 2013), where most beneficial mutations detected at high salinity were deleterious in other environments. A potential explanation for this discrepancy lies in the choice of the protein region. Whereas the previous study focused on only nine amino acids and thus chose a set of positions that are likely to show functionally interesting patterns, we here had the experimental power to scan a larger region of the protein. Along a larger, less specifically chosen, stretch of the protein, it is likely that many positions and mutations play the same role across many environments. Consequently, only a small number of mutations in this greater set may have a functionally different role across environments, leading to costs of adaptation. In such cases, costs of adaptation of beneficial mutations could still be large, but due to their low proportion they barely affect the correlation of selection coefficients across environments.

255 **Beneficial mutations appear across all environments and are enriched in a region of**
256 **Hsp90 that is implicated in stabilization of interdomains and client binding interfaces.**
257 Next, we identified the proportion and identity of putatively beneficial mutations across all
258 treatments (Supplementary Table 1, Supplementary Figure 4, 5). We categorized mutants as
259 putatively beneficial if the lower limit of the 95% confidence interval of the selection
260 coefficient was larger than 0, and we considered as strongest candidates those that overlapped
261 between replicates from the same environment. As reported above, wild-type like mutations
262 are the most abundant category in all environments, with a low proportion but considerable
263 number of beneficial mutations in almost all environments (Supplementary Figure 4, 5). We
264 find the lowest number of beneficial mutations in H_2O_2 and salt environments (n_{H2O2}=65 and
265 n_{Salt}=64, around 2.5% of the mutations) and the highest in diamide ($n_{Diamide}$=307, 12.3%).
266 Interestingly, in diamide beneficial mutations were dispersed across the whole middle
267 domain (Supplementary Figure 4), whereas most of the other environments showed a clear
268 enrichment of beneficial mutations in the region of positions 356-366 (Figure 3,
269 Supplementary Figure 4). This hotspot of beneficial mutations was particularly strong in
270 sorbitol (Supplementary Figure 4). Out of the 31 beneficial mutations that show a beneficial
271 effect in at least two environments, 18 are located in this hotspot.

272

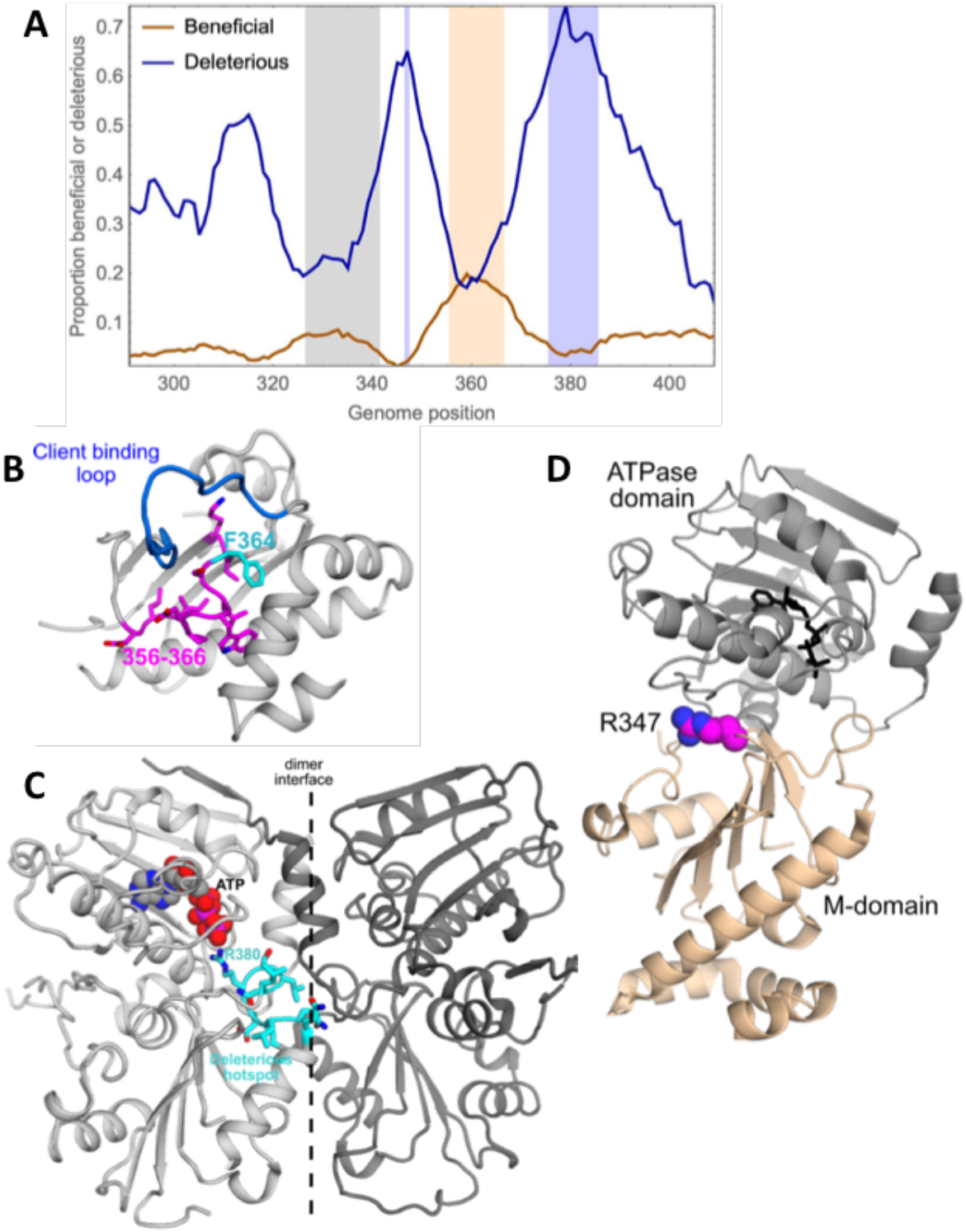


Figure 3: **Hotspots of beneficial and deleterious mutations in the middle domain do not coincide with known client binding loop.** A) Average proportion of beneficial and deleterious mutations along the genome across all environments, illustrated using a 10-amino-acid sliding window. The region with the 10% largest proportions of beneficials is highlighted in light orange (position 356-366), the region with the 10% largest proportions of deleterious mutations is highlighted in light blue (positions 347 and 376-385). A known

client binding loop (positions 327-341) is highlighted in light gray. To avoid biases, the analysis was restricted to subsets of two replicates of all normal-expression environments. B) The beneficial hotspot is adjacent to a known client binding loop. Structural representation of the Hsp90 middle domain (1HK7.PDB) illustrating the beneficial hotspot in magenta with residue 364 highlighted in cyan. Biochemical, structural and mutational studies identified positions 327-341 as a client binding loop, which is shown in blue. C) The deleterious hotspot contains a catalytic amino acid and is at a dimerization interface. Structural representation of the ATPase and middle domains of Hsp90 from 2CG9.PDB. The deleterious hotspot is colored in cyan and includes R380 that stabilizes the leaving phosphate during ATP hydrolysis. Subunits are distinguished with different shading. D) The deleterious hotspot at position 347 is also located at the interface of the ATPase domain and the middle domain, and is highlighted in pink and blue.

Interestingly, the beneficial hotspot contains residues that are part of an allosteric center which mediates distinct conformations in structures with and without client (Czemeres et al. 2017). Despite its different structural arrangements, the hydrophobic amino acids in this region are mostly buried from solvent (Figure 3B), indicating that the primary role of this region is to mediate the stability and rearrangement of different conformations. Of note, the position with the greatest number of beneficial mutations is F364 (Supplementary Figure 6). The burial of the large hydrophobic side chain of F364 should provide local stability. The beneficial mutations at position 364 and in this region in general suggest that disruption of local stability and conformational dynamics may alter Hsp90 function. Based on the prominent role of the middle domain in binding to clients, we hypothesize that Hsp90 function may be altered by the identified candidate beneficial mutations in a manner that changes the relative affinity and thereby priority of different clients - a property that could provide benefits in specific conditions. Consistent with this hypothesis, the beneficial hotspot is adjacent to a known client binding loop (Figure 3) such that alterations in the structure or dynamics of the hotspot are likely to influence client binding nearby. It is possible that natural selection balances client priorities integrated over multiple conditions, which provides opportunities for Hsp90 mutations to improve priorities for individual conditions.

311 **The proportion of deleterious, wild-type like, and beneficial mutations varies greatly**
312 **along the protein sequence.** We showed above that the correlations of fitness effects across
313 environments are generally large, which indicates similar effects of the same mutations
314 across environments. In contrast, the proportions of beneficial, wild-type like, and deleterious
315 mutations vary greatly along the protein sequence. Whereas the overall pattern is similar
316 between environments, the relative proportions differ between environments (Supplementary
317 Figure 4, 5). For example, at low expression there are less beneficial mutations, but the
318 regions that show deleterious mutations are similar between expression levels
319 (Supplementary Figure 4). Our finding that the same positions are enriched for deleterious
320 and beneficial mutations across environments suggests that the structural properties of the
321 middle domain in these regions, rather than its binding partners, might be the most important
322 factor for predicting the fitness effects of mutations.
323
324 One position that particularly stands out with respect to the proportions of beneficial, wild-
325 type like and deleterious mutations is 364. Position 364 lies at the center of the cluster of the
326 10% largest average proportions of beneficial (Figure 3B) and shows a clear deviation from
327 the usually observed DFE with an enrichment of beneficials (Supplementary Figure 6).
328 Specifically, we identified 6 mutations at this position that are beneficial in at least two
329 environments (Supplementary Figure 6) . As mentioned above this may be due to a disruption
330 of the stability of the protein that may alter the relative affinity of Hsp90 to its clients, which
331 may be beneficial in specific environments.
332
333 Interestingly, the regions with the largest proportions of deleterious mutations are located
334 near the ends of the beneficial hotspot region (position 347 and 376-385, respectively)
335 (Supplementary Figure 4). At the latter deleterious hotspot, around 70% of all mutations are
336 deleterious in a sliding window of 10 amino-acid positions. Such a shift of the DFE towards
337 deleterious mutations suggests that this protein region is under strong purifying selection.
338 Structurally, the region 376-385 is part of a catalytic loop required for ATP hydrolysis which
339 is necessary for the activation of all known clients (Wolmarans et al. 2016). Position R380 in
340 the catalytic loop binds to and stabilizes the leaving phosphate of ATP (Figure 3C) and
341 mutations at this position compromise Hsp90 function and cell viability (Meyer et al.2003).
342 The efficiency of catalysts depends strongly on geometry, and the precise location of R380
343 relative to ATP is likely tightly linked to Hsp90 function. The regions adjacent to R380

344 appear to be important for positioning the catalytic arginine, providing a rationale for the
345 strong purifying selection that we observe. The other deleterious hotspot at position 347 is
346 also located at the interface of the ATPase domain and the middle domain (Figure 3D),
347 consistent with current understanding that Hsp90 mechanism requires precise ATP-dependent
348 interactions between these domains (Schopf et al. 2017).

349 **Few mutations show costs of adaptation.** Among the identified beneficial mutations, we
350 were specifically interested in those that are deleterious in other environments, which results
351 in a so called "cost of adaptation". The same phenomenon has also been termed antagonistic
352 pleiotropy. In previous work we reported a large prevalence of costs of adaptation in a 9-
353 amino-acid region of Hsp90 (Hietpas et al 2013). In this study, we found that costs of
354 adaptation are not pervasive (Figure 4, Supplementary Table 3). Consistent with the
355 comparatively low correlations of fitness effects and the special role of the diamide stress
356 discussed above, we observed the largest proportion of mutations that show costs of
357 adaptation between diamide and other environments. Whereas the mean proportion of
358 mutations displaying costs of adaptation across all comparisons of environments was 14.3%,
359 comparisons that included diamide showed on average 22.7% mutations with costs of
360 adaptation. The largest proportions of mutations with costs of adaptation (>40%) were seen
361 in comparisons of low-expression diamide and high-expression H_2O_2.

362

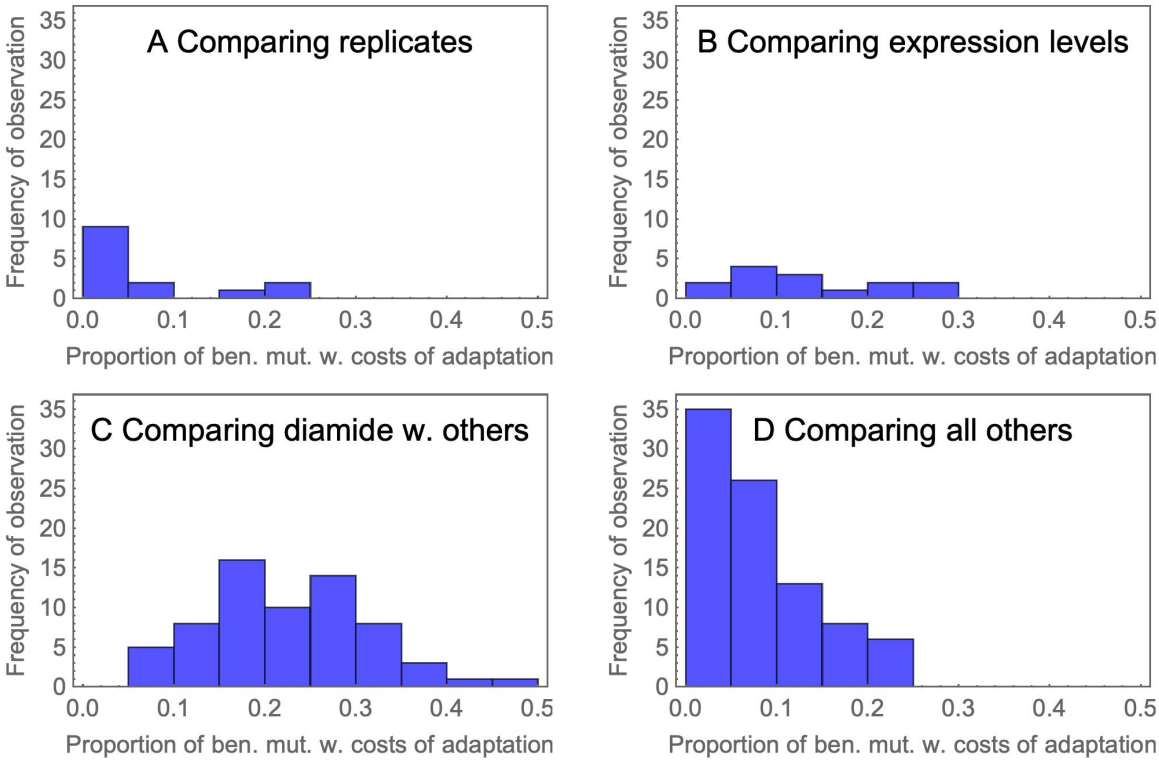


Figure 4: Histogram of proportions of mutations that display costs of adaptation illustrates that costs of adaptation are generally rare, and most likely if comparisons involve diamide environments. The histograms correspond to comparisons of replicates (A), expression levels (B), the subset of other comparisons that include diamide (C), and all other comparisons (D). There is an enrichment of mutations that show costs of adaptation in comparisons between all environments and diamide, compared to comparisons between replicates, expression levels and other environments. This is consistent with what was observed above and further supports that the stress through exposure to diamide is different from all other environments.

Mapping the proportion of beneficials that are deleterious in other environments along the protein sequence, we observed that there is a hotspot for costs of adaptation at amino-acid positions 381-391 (Figure 5). Interestingly, this region showed costs of adaptation across various environments, which indicates that each environment has specific beneficial mutations that are deleterious in other environments. Structurally, this region also belongs to the catalytic loop involved in ATP hydrolysis discussed above. Indeed, the identified region partly overlaps with the above-discussed hotspot of deleterious mutations (positions 376-385), which is unsurprising since a larger proportion of deleterious mutations also increases the (technical) proportion of mutations that can be classified as costly.

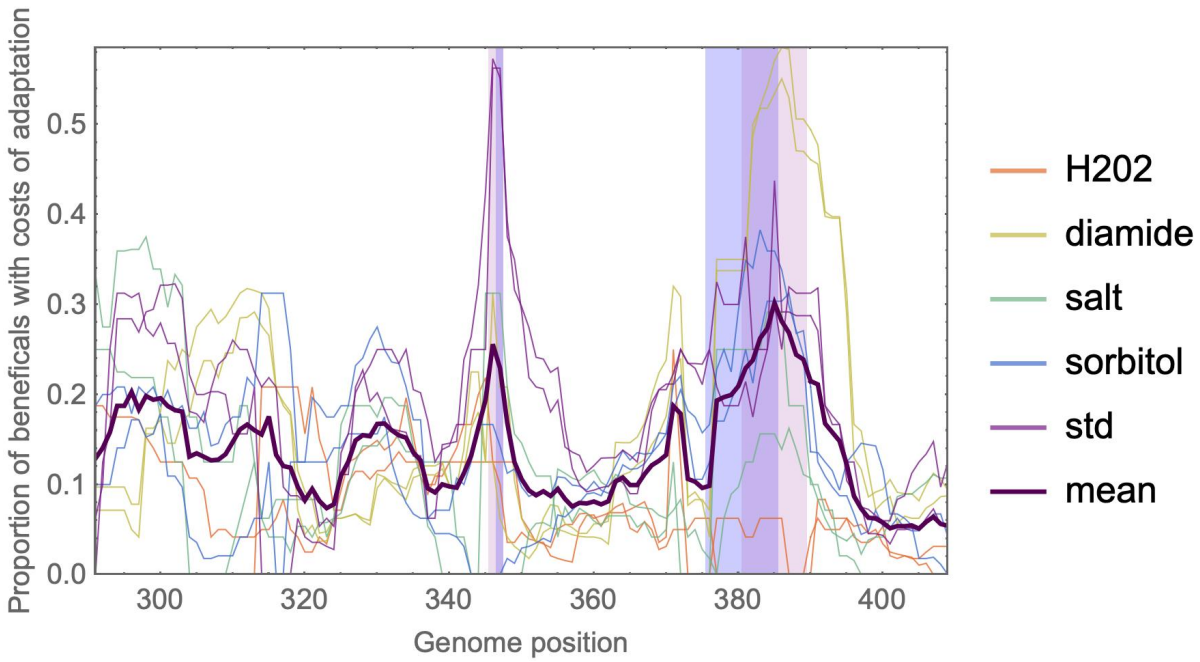


Figure 5: A hotspot of costs of adaptation is located at positions 381-391. The 10% positions with the largest mean proportion of mutations that display costs of adaptation (dark purple line) are highlighted in light purple. This region greatly overlaps with the hotspots of deleterious mutations from Figure 3 (highlighted in blue). Thin curves indicate the mean proportion of beneficial mutations in a focal environment that is deleterious in another environment. A 10-amino-acid sliding window was used to locate region-specific effects. To avoid biases, the analysis was restricted to subsets of two replicates of all normal-expression environments.

We next computed whether there was a correlation between the effect of beneficial mutations in one environment with their effect in the other environments. Indeed, the effects of beneficial mutations in diamide are negatively correlated with the same mutation's effect in all other environments (Figure 5, Supplementary Figure 7) except H_20_2. In other words, the more beneficial a mutation was in diamide, the more deleterious tended to be in salt, standard and sorbitol. Again, this suggests that the beneficial mutations we identified in diamide may be involved in the response to a very specific type of stress. Across all other environments, we did not see any evidence that stronger beneficial mutations tended to have a more deleterious effect in other environments.

405 **Conclusion**

406 Recent advances in experimental and technological approaches have led to the feasibility of
407 large-scale screens of the DFE of new mutations, which, from an evolutionary point of view,
408 is a key entity to determine the potential for adaptation. Here, we took a new route by
409 mapping the proportions of beneficial, wild-type like, and deleterious mutations along a 119-
410 amino-acid region of the Hsp90 protein in yeast, a protein that is heavily involved in the
411 response to environmental stressors. By comparing the DFE along the protein sequence and
412 between environments, we identified hotspots of beneficial and deleterious mutations that are
413 common across environments, and a region in which beneficial mutations in one environment
414 tend to be deleterious in other environments. Interestingly, neither of these regions coincided
415 with the best described client binding loop in the studied region, which we had *a priori*
416 considered the most likely candidate to display patterns different from the rest of the region.
417 Moreover, our analyses suggested that mutational effects generally differ little across
418 environments except in diamide, which stands out both with respect to the number and also
419 the distribution of beneficial mutations along the studied protein region. Altogether, our study
420 of the DFE across environments sheds light on the evolutionary role of a specific protein
421 region from a new perspective.

422 **Materials and Methods**

423 **Generating point mutations**. To accurately measure the fitness effects of all possible point
424 mutations in a large portion of the middle (client-binding) domain of yeast Hsp90, we used
425 saturation mutagenesis at positions 291-409 (see Supplementary figure 8). We used a cassette
426 ligation strategy to introduce mutations as previously described (Hietpas et al. 2012). This
427 strategy reduces the likelihood of secondary mutations because it avoids the potential for
428 errors during PCR steps. As a control for the mutational procedure, twelve positions were
429 randomized in isolation and Sanger sequenced to assess the level of incorporation of all four
430 nucleotides at each position in the target codon. At all randomized nucleotide positions
431 within these twelve samples, we observed a similar magnitude of signal for each of the four
432 nucleotides. For larger scale production, libraries were generated at 10 positions at a time as
433 previously described (Hietpas et al. 2012). As additional controls, we generated a sample
434 containing individual stop codons as well as the parental wildtype Hsp90 sequence. All
435 variants of Hsp90 were generated in a plasmid (pRS414GPD) previously shown to produce
436 endogenous levels of Hsp90 protein in yeast (Chang and Lindquist 1994).

437

438 **Constructing barcoded libraries.** To improve the efficiency and accuracy of fitness
439 estimates, we added barcodes to a non-functional region of the plasmid, ~200 NTs
440 downstream from the 3' untranslated region of Hsp90. Barcodes were introduced using a
441 cassette ligation strategy (Supplementary Figure 8). Plasmid libraries were treated with NotI
442 and AscI to generate directional sticky ends. NotI and AscI recognize 8-base cut sites that are
443 unlikely to cut any of the Hsp90 variants in the library. We designed and annealed barcode
444 forward and barcode reverse oligos together such that the resulting duplex product included a
445 central N18 region bracketed by constant regions that facilitate annealing and overhangs that
446 direct directional ligation into NotI and AscI overhangs. One of the constant ends in the
447 designed oligo cassette contains an annealing region for an Illumina sequencing primer.
448 Barcoded libraries were transformed into *E. coli* and pooled into a bulk culture that contained
449 about 10-fold more transformants than Hsp90 variants in the library. We purified barcoded
450 plasmids from this bulk culture. This procedure resulted in approximately 10 barcodes for
451 each Hsp90 codon variant in the library (Supplementary Figure 9). The potential diversity in
452 the N18 barcode that we used ($4^{18} \sim 10^{11}$) far exceeds the number of barcodes that we utilize
453 (~64*119*10~10^5), which makes it likely that each Hsp90 variant will have a barcode that
454 differs from all other barcodes at multiple nucleotides. With this setup, errors in sequencing
455 of barcodes can be detected and eliminated from further analysis, which reduces the impact
456 of sequencing misreads on estimates of variant frequency and fitness. Additional controls
457 consisting of individual stop codons and wildtype Hsp90 were barcoded separately. For these
458 controls, we isolated barcoded plasmid DNA from individual bacterial colonies and
459 determined the barcodes by Sanger sequencing (Supplementary Figure 8).

460 **Associating barcodes to mutants.** To identify the barcodes associated with each Hsp90
461 variant in our libraries, we used a paired-end sequencing approach essentially as previously
462 described (Hiatt et al. 2010, see also Supplementary Figure 10). Using paired-end
463 sequencing on an Illumina MiSeq Instrument barcodes were associated with variant
464 genotypes via reading from the Hsp90 gene with a 250 base-pair read and the associated N18
465 barcode with a 50 base-pair read. To reduce the size of the DNA fragments for efficient
466 Illumina sequencing, we removed a portion of the Hsp90 gene such that the randomized
467 regions were closer to the N18 barcode. This was done to increase the density of DNA on the
468 sequencer by reducing the radius of clonal clusters during sequencing. Plasmid DNA was
469 linearized with StuI and NotI endonucleases that removed ~400 bp's of DNA. The linearized

470 products were circularized by blunt ending followed by ligation at DNA concentrations that
471 predominantly lead to unimolecular ligations (Revie et al 1998) . The resulting DNA
472 products were amplified using a PE2_F primer and the standard Illumina PE1 primer that
473 anneals next to the N18 barcode. Two PE2_F primers were designed in order to read across
474 the region of Hsp90 that we randomized. PCR products were gel purified and submitted for
475 paired end sequencing. We obtained sufficient paired-end reads such that the average barcode
476 was read more than 10 times. The paired-end sequencing data was subjected to a custom data
477 analysis pipeline to associate Hsp90 variants with barcodes. First, very low-quality reads with
478 any Phred score less than 10 in reads 1 or 2 were discarded. Next, the data were organized by
479 the barcode sequence. For barcodes with three or more reads, we constructed a consensus of
480 the Hsp90 sequence read. We compared the consensus sequence to the parental Hsp90
481 sequence in order to determine mutations. Consensus sequences containing more than one
482 protein mutation were discarded. The pipeline output generated a file organized by point
483 mutation that lists all barcodes associated with that mutation (Supplementary Data). This file
484 was used as the basis for calling a variant based on barcode reads.

485 **Yeast competitions.** As in previous work (Hietpas et al. 2012, Hietpas et al. 2013), we
486 performed Hsp90 competitions using a shutoff strain of *S. cerevisiae* (DBY288) (
487 Supplementary figure 11). The sole copy of Hsp90 in DBY288 is driven by a galactose-
488 inducible promoter, such that the strain requires galactose for viability and cannot grow on
489 glucose. The introduction of a functional Hsp90 variant driven by a constitutive promoter
490 rescues the growth of DBY288 in glucose media. We introduced the library of middle
491 domain variants of Hsp90 driven by a constitutive promoter into DBY288 cells using the
492 lithium acetate method (Gietz and Schiestl 2007). The efficiency of transformation was more
493 than 10-fold higher than the number of barcodes, such that the average barcode was
494 transformed into more than 10 individual cells. To enable the analyses of full biological
495 replicates, transformations were performed in replicate such that a separate population of
496 yeast were transformed for each biological replicate. Transformed yeast were initially
497 selected in RGal-W (synthetic media that lacked tryptophan and contained 1% raffinose and
498 1% galactose supplemented with 100 μg/mL ampicillin to hinder bacterial contamination) to
499 select for the plasmid, but not function of the plasmid encoded Hsp90 variant. This enabled
500 us to generate a yeast library containing Hsp90 variants that could support a full range of
501 fitness from null to adaptive. Cells were grown for 48 hours at 30 °C in liquid SRGal-W until

502 the culture was visibly opaque compared to a control transformation sample that lacked
503 plasmid DNA but was otherwise identical to the library sample.

504 To initiate the competition, cells were transferred to shut-off conditions and different
505 environmental stresses. In order to deplete the pool of wild type Hsp90 protein, the library
506 sample was diluted into SD-W (synthetic media lacking tryptophan with 2% glucose and 100
507 µg/mL of ampicillin) and grown at 30 °C for 10 hours. After depletion of wild type Hsp90
508 protein, the cells were split into different stress conditions in SD-W media at 30 °C: salt
509 stress (0.5M NaCl), osmotic stress (0.6M sorbitol), oxidative stress (0.6 mM hydrogen
510 peroxide or 0.85 mM diamide), as well as a control non-stress condition. We used a custom
511 built turbidostat (Supplementary figure 11) to provide constant growth conditions during this
512 competition phase of the experiment. Cells were grown under consistent density and
513 population size. 10^9 cells were maintained in a 50 mL volume over the course of 40 hours for
514 each condition. Rapid magnetic stirring was used to provide aeration to the media and cells.
515 Samples of 10^8 cells were collected after 0, 4, 8, 12, 24, 32, 40, and 48 hours of competition
516 in the different conditions. At the time of collection, samples were centrifuged and the pellets
517 immediately frozen at -80 °C.

518

519

520

521 **Sequencing of competition samples.** The fitness effects of Hsp90 variants were estimated
522 based on frequencies observed in next-generation sequencing analyses essentially as
523 previously described (Hietpas et al. 2012). DNA was isolated from each timepoint sample as
524 described (Hietpas et al. 2013) and the barcodes were amplified and sequenced. Barcodes
525 were amplified using the standard Illumina PE1 primer that anneals next to the N18 barcode
526 and custom designed barcode_forward primers. A set of barcode_forward primers were
527 designed with identifier sequences that could be read during sequencing and used to
528 distinguish each timepoint sample. Each identifier sequence was eight nucleotides in length
529 and differed by at least two bases from all other identifier sequences. Twenty cycles of PCR
530 were sufficient to generate clear products from all samples. These PCR products were
531 purified on silica columns (Zymo Research) and sequenced for 100 base single reads on an
532 Illumina NextSeq instrument. The barcode corresponded to the first 18 bases and the
533 identifier at positions 91-98. The resulting fastq files were processed and analyzed using
534 customized software tools. First, poor quality reads containing any positions with a Phred

535 score less than 20 were discarded. Reads were tabulated if the barcode matched a barcode
536 associated with a point mutation and if the identifier matched with a timepoint. The analyses
537 scripts output a file with the number of reads for each amino acid point mutant at each
538 timepoint in each condition (Supplementary Data). Barcodes with 0 reads at the first time
539 point, or with a total of 1 read along the whole trajectory were removed from the analysis.

540 **Estimation of selection coefficients.** Inference of selection coefficients was performed via
541 log-linear regression as described in Matuszewski et al 2016 (see also Supplementary
542 Material). To improve estimation accuracy by incorporating information from each individual
543 barcode, the linear model for each amino-acid changing mutation included barcode identities
544 as nominal variables. For each amino-acid changing mutation we obtained a selection
545 coefficient and 95% confidence interval (CI) of the estimate, representing the variation within
546 amino acid due to differences in codon*Barcode tag throughout time. This reduces the impact
547 of potential outliers and averages over synonymous codons within amino acid. Mutants with
548 50 or less total reads at the first time point were removed from the data set. Finally, we
549 normalized all selection coefficients, by subtracting the median of all mutations that were
550 synonymous to the wild type. This ensures that the average of the selection coefficients of
551 wild type synonyms represents a selection coefficient of 0. To categorize mutations as
552 beneficial, deleterious or neutral we tested the overlap of the CI with 0. Namely, if the lower
553 CI limit was larger than 0, a mutation was considered beneficial, and if the upper CI limit was
554 smaller than 0, a mutation was considered deleterious. All other mutations were considered
555 wild-type like.
556

557 **Costs of adaptation.** To compute the costs of adaptation for Figure 4 and 5, we first selected
558 all mutations in a focal environment that were categorized as beneficial. We then computed
559 the proportion of this subset of mutations that was deleterious either in one other environment
560 (for Figure 4) or across all other environments (excluding replicates of the focal environment
561 and low-expression environments; for Figure 5). Calculating the proportion of mutations with
562 costs of adaptation compensates for variable numbers of beneficial mutations across
563 environments, which may occur due to both biological and experimental reasons (e.g.,
564 different error margins). For Figure 2 and 4, all available data sets were considered for the
565 analysis. For Figure 3 and 5, only two replicates of each environment were considered, and
566 the low-expression data were excluded. That was in order to avoid biases due to the under- or
567 overrepresentation of environments in the analysis.

Analyses were performed using R version 3.5.1 (R Core Team 2018), Mathematica 12.0.0.0, Python 2.7.12 and Pymol 0.99. Code is available as Supplementary Material.
All data will be made available in a data repository upon acceptance of the manuscript.

Acknowledgements

This work was supported in part by grant R01GM112844 from the National Institutes of Health (to DNAB). IF was supported by a postdoctoral fellowship from the FCT (Fundação para a Ciência e a Tecnologia) within the project JPIAMR/0001/2016. CB is grateful for support by EMBO Installation Grant IG4152 and by ERC Starting Grant 804569 - FIT2GO.

References

Acevedo, A., Brodsky, L., & Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature, 505(7485), 686–690.

Arribas, M., K. Kubota, L. Cabanillas, and E. Lazaro. 2014. Adaptation to fluctuating temperatures in an RNA virus is driven by the most stringent selective pressure. PLoS One 9:e100940.

Bank, C., Hietpas, R.T., Wong, A., Bolon, D.N. and Jensen, J.D., 2014 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 196(3): 841–852.

Barrett, R. D. and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38-44.

Bataillon, T., Zhang, T., & Kassen, R. (2011). Cost of adaptation and fitness effects of beneficial mutations in Pseudomonas fluorescens. Genetics, 189(3), 939–949.

Bataillon T., Bailey S.F. 2014 Effects of new mutations on fitness: insights from models and data. Ann N Y Acad Sci. 1320:76-92.

Borkovich, K. A., F. W. Farrelly, D. B. Finkelstein, J. Taulien, and S. Lindquist. 1989. hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. Mol. Cell. Biol. 9:3919-3930.

598 Boucher, J. I., Cote, P., Flynn, J., Jiang, L., Laban, A., Mishra, P., Roscoe B.P. & Bolon, D.
599 N. (2014). Viewing protein fitness landscapes through a next-gen lens. Genetics, 198(2),
600 461–471.
601 Boucher, J. I., Bolon, D. N., & Tawfik, D. S. (2016). Quantifying and understanding the
602 fitness effects of protein mutations: Laboratory versus nature. Protein Science, 25(7), 1219-
603 1226.
604 Brennan, G. L., N. Colegrave, and S. Collins. 2017. Evolutionary consequences of
605 multidriver environmental change in an aquatic primary producer. Proc. Natl. Acad. Sci. U.
606 S. A. 114:9930-9935.
607 Chang, H. C. and S. Lindquist. 1994. Conservation of Hsp90 macromolecular complexes in
608 Saccharomyces cerevisiae. J. Biol. Chem. 269:24983-24988.
609 Chen, B., M. E. Feder, and L. Kang. 2018. Evolution of heat-shock protein expression
610 underlying adaptive responses to environmental stress. Mol. Ecol. 27:3040-3054.
611 Cunningham, C. N., K. A. Krukenberg, and D. A. Agard. 2008. Intra- and intermonomer
612 interactions are required to synergistically facilitate ATP hydrolysis in Hsp90. J. Biol. Chem.
613 283:21170-21178.
614 Czemeres, J., K. Buse, and G. M. Verkhivker. 2017. Atomistic simulations and network-
615 based modeling of the Hsp90-Cdc37 chaperone binding with Cdk4 client protein: A
616 mechanism of chaperoning kinase clients by exploiting weak spots of intrinsically dynamic
617 kinase domains. PLoS One 12:e0190267.
618 de Nadal, E., G. Ammerer, and F. Posas. 2011. Controlling gene expression in response to
619 stress. Nature reviews. Genetics 12:833-845.
620 Dekel, E. and U. Alon. 2005. Optimality and evolutionary tuning of the expression level of a
621 protein. Nature 436:588-592.
622 Dhar, R., R. Sagesser, C. Weikert, J. Yuan, and A. Wagner. 2011. Adaptation of
623 Saccharomyces cerevisiae to saline stress through laboratory evolution. J. Evol. Biol.
624 24:1135-1153.
625 Doud, M. and Bloom, J., 2016 Accurate Measurement of the Effects of All Amino-Acid
626 Mutations on Influenza Hemagglutinin, Viruses 8(6), 155.
627 Eyre-Walker, A. and Keightley, P.D. 2007 The distribution of fitness effects of new
628 mutations. Nature Review Genetics 8(8): 610.
629 Firnberg E, Labonte JW, Gray JJ, Ostermeier M. 2014 A Comprehensive, High-Resolution
630 Map of a Gene's Fitness Landscape. Mol Biol Evol..

631 Fitzgerald, D. M. and S. M. Rosenberg. 2019. What is mutation? A chapter in the series: How
632 microbes "jeopardize" the modern synthesis. PLoS Genet. 15:e1007995.
633 Flynn, J. M., Rossouw, A., Cote-Hammarlof, P. A., Fragata, I., Mavor, D., Hollins, C., Bank,
634 C. & Bolon, D. N. (2019). Investigating the influence of environment on the evolution of
635 Hsp90 using comprehensive fitness maps. bioRxiv, 823468.
636 Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., &
637 Fields, S. (2010). High-resolution mapping of protein sequence-function relationships. Nature
638 methods, 7(9), 741–746.
639 Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D.
640 Botstein, and P. O. Brown. 2000. Genomic expression programs in the response of yeast cells
641 to environmental changes. Mol. Biol. Cell 11:4241-4257.
642 Genest, O., M. Reidy, T. O. Street, J. R. Hoskins, J. L. Camberg, D. A. Agard, D. C.
643 Masison, and S. Wickner. 2013. Uncovering a region of heat shock protein 90 important for
644 client binding in E. coli and chaperone function in yeast. Mol. Cell 49:464-473.
645 Gietz, R. D. and R. H. Schiestl. 2007. High-efficiency yeast transformation using the
646 LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2:31-34.
647 Gopinath, R. K., S. T. You, K. Y. Chien, K. B. Swamy, J. S. Yu, S. C. Schuyler, and J. Y.
648 Leu. 2014. The Hsp90-dependent proteome is conserved and enriched for hub proteins with
649 high levels of protein-protein connectivity. Genome Biol. Evol. 6:2851-2865.
650 Hagn, F., S. Lagleder, M. Retzlaff, J. Rohrberg, O. Demmer, K. Richter, J. Buchner, and H.
651 Kessler. 2011. Structural analysis of the interaction between Hsp90 and the tumor suppressor
652 protein p53. Nat. Struct. Mol. Biol. 18:1086-1093.
653 Hawle, P., M. Siepmann, A. Harst, M. Siderius, H. P. Reusch, and W. M. Obermann. 2006.
654 The middle domain of Hsp90 acts as a discriminator between different types of client
655 proteins. Mol. Cell. Biol. 26:8385-8395.
656 Hiatt, J. B., R. P. Patwardhan, E. H. Turner, C. Lee, and J. Shendure. 2010. Parallel, tag-
657 directed assembly of locally derived short sequence reads. Nature methods 7:119-122.
658 Hietpas, R. T., Jensen, J. D., & Bolon, D. N. (2011). Experimental illumination of a fitness
659 landscape. Proceedings of the National Academy of Sciences, 108(19), 7896-7901.
660 Hietpas, R., B. Roscoe, L. Jiang, and D. N. Bolon. 2012. Fitness analyses of all possible point
661 mutations for regions of genes in yeast. Nat. Protoc. 7:1382-1396.
662 Hietpas, R. T., C. Bank, J. D. Jensen, and D. N. A. Bolon. 2013. Shifting fitness landscapes
663 in response to altered environments. Evolution 67:3512-3522.

664 Hom N., Gentles L., Bloom J.D., and Lee K.K. 2019 Deep Mutational Scan of the Highly
665 Conserved Influenza A Virus M1 Matrix Protein Reveals Substantial Intrinsic Mutational
666 Tolerance. Journal of Virology 93(13):e00161-19.
667 Jarosz, D. F. and S. Lindquist. 2010. Hsp90 and environmental stress transform the adaptive
668 value of natural genetic variation. Science 330:1820-1824.
669 Jiang, L., Mishra, P., Hietpas, R. T., Zeldovich, K. B., & Bolon, D. N. (2013). Latent effects
670 of Hsp90 mutants revealed at reduced expression levels. PLoS genetics, 9(6), e1003600.
671 Kaplan, K. B. and R. Li. 2012. A prescription for 'stress'--the role of Hsp90 in genome
672 stability and cellular adaptation. Trends Cell Biol. 22:576-583.
673 Karagoz, G. E., A. M. Duarte, E. Akoury, H. Ippel, J. Biernat, T. Moran Luengo, M. Radli, T.
674 Didenko, B. A. Nordhues, D. B. Veprintsev, C. A. Dickey, E. Mandelkow, M. Zweckstetter,
675 R. Boelens, T. Madl, and S. G. Rudiger. 2014. Hsp90-Tau complex reveals molecular basis
676 for specificity in chaperone action. Cell 156:963-974.
677 Karagoz, G. E. and S. G. Rudiger. 2015. Hsp90 interaction with clients. Trends Biochem.
678 Sci. 40:117-125.
679 Kvitek DJ, Sherlock G 2013 Whole Genome, Whole Population Sequencing Reveals That
680 Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment.
681 PLOS Genetics 9(11): e1003972
682 Lindquist, S. 2009. Protein folding sculpting evolutionary change. Cold Spring Harb. Symp.
683 Quant. Biol. 74:103-108.
684 Loewe, L. 2010 The population genetics of mutations: good, bad and indifferent.
685 Philosophical Transactions of the Royal Society B: Biological Sciences 365: 1153-1167.
686 Manceau, M., V. S. Domingues, R. Mallarino, and H. E. Hoekstra. 2011. The developmental
687 role of Agouti in color pattern evolution. Science 331:1062-1065.
688 Matuszewski, S., Hildebrandt, M. E., Ghenu,A.-H., Jensen, J. D., and Bank, C. 2016 A
689 Statistical Guide to the Design of Deep Mutational Scanning Experiments, Genetics 204(1),
690 77–87.
691 McClellan, A. J., Y. Xia, A. M. Deutschbauer, R. W. Davis, M. Gerstein, and J. Frydman.
692 2007. Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems
693 approaches. Cell 131:121-135.
694 Melamed D., Young D. L., Gamble C. E., Miller C. R., Fields S., 2013. Deep mutational
695 scanning of an RRM domain of the Saccharomyces cerevisiae poly(A)-binding protein. RNA
696 19: 1537–1551.

Meyer, P., C. Prodromou, B. Hu, C. Vaughan, S. M. Roe, B. Panaretou, P. W. Piper, and L. H. Pearl. 2003. Structural and functional analysis of the middle segment of hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. Mol. Cell 11:647-658.

Mumby, P. J. and R. van Woesik. 2014. Consequences of ecological, evolutionary and biogeochemical uncertainty for coral reef responses to climatic stress. Curr. Biol. 24:R413-423.

Mustonen V and Lässig M. 2009 From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet. 25(3):111-9.

Nathan, D. F. and S. Lindquist. 1995. Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase. Mol. Cell. Biol. 15:3917-3925.

Nathan, D. F., M. H. Vos, and S. Lindquist. 1997. In vivo functions of the Saccharomyces cerevisiae Hsp90 chaperone. Proc. Natl. Acad. Sci. U. S. A. 94:12949-12956.

Ohta T. 1992 The Nearly Neutral Theory of Molecular Evolution. Annual Review of Ecology and Systematics 23(1):263-286.

Park, S. J., B. N. Borin, M. A. Martinez-Yamout, and H. J. Dyson. 2011a. The client protein p53 adopts a molten globule-like state in the presence of Hsp90. Nat. Struct. Mol. Biol. 18:537-541.

Park, S. J., M. Kostic, and H. J. Dyson. 2011b. Dynamic Interaction of Hsp90 with Its Client Protein p53. J. Mol. Biol. 411:158-173.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reznick, D.N. and Ghalambor, C.K. 2005 Selection in Nature: Experimental Manipulations of Natural Populations, Integrative and Comparative Biology, 45(3):456–462.

Revie, D., Smith, D. W., & Yee, T. W. (1988). Kinetic analysis for optimization of DNA ligation reactions. Nucleic acids research, 16(21), 10301-10321.

Richter, K., M. Haslbeck, and J. Buchner. 2010. The heat shock response: life on the verge of death. Mol. Cell 40:253-266.

Rohl, A., J. Rohrberg, and J. Buchner. 2013. The chaperone Hsp90: changing partners for demanding clients. Trends Biochem. Sci. 38:253-262.

Rutherford, S., Y. Hirate, and B. J. Swalla. 2007. The Hsp90 capacitor, developmental remodeling, and evolution: the robustness of gene networks and the curious evolvability of metamorphosis. Crit. Rev. Biochem. Mol. Biol. 42:355-372.

730 Rutherford, S. L. 2003. Between genotype and phenotype: protein chaperones and
731 evolvability. Nature reviews. Genetics 4:263-274.
732 Sane, M. , Miranda, J. J. and Agashe, D. (2018), Antagonistic pleiotropy for carbon use is
733 rare in new mutations. Evolution, 72: 2202-2213.
734 Sangster, T. A., S. Lindquist, and C. Queitsch. 2004. Under cover: causes, effects and
735 implications of Hsp90-mediated genetic capacitance. Bioessays 26:348-362.
736 Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Ivankov
737 DN, Bozhanova NG, Baranov MS, Soylemez O, Bogatyreva NS, Vlasov PK, Egorov ES,
738 Logacheva MD, Kondrashov AS, Chudakov DM, Putintseva EV, Mamedov IZ, Tawfik DS,
739 Lukyanov KA, Kondrashov FA. 2016 Local fitness landscape of the green fluorescent
740 protein. Nature 533(7603):397-401.
741 Sato, S., N. Fujita, and T. Tsuruo. 2000. Modulation of Akt kinase activity by binding to
742 Hsp90. Proc. Natl. Acad. Sci. U. S. A. 97:10832-10837.
743 Schopf, F. H., Biebl, M. M., & Buchner, J. 2017. The HSP90 chaperone machinery. Nature
744 reviews Molecular cell biology, 18(6), 345.
745 Sniegowski P.D. and Gerrish P.J. 2010 Beneficial mutations and the dynamics of adaptation
746 in asexual populations. Philos Trans R Soc Lond B Biol Sci. 365(1544):1255-63.
747 Soskine M, Tawfik DS. 2010 Mutational effects and the evolution of new protein functions.
748 Nat Rev Genet; 11(8):572–82
749 Street, T. O., L. A. Lavery, K. A. Verba, C. T. Lee, M. P. Mayer, and D. A. Agard. 2012.
750 Cross-monomer substrate contacts reposition the Hsp90 N-terminal domain and prime the
751 chaperone activity. J. Mol. Biol. 415:3-15.
752 Taipale, M., D. F. Jarosz, and S. Lindquist. 2010. HSP90 at the hub of protein homeostasis:
753 emerging mechanistic insights. Nat. Rev. Mol. Cell Biol. 11:515-528.
754 Taipale, M., I. Krykbaeva, M. Koeva, C. Kayatekin, K. D. Westover, G. I. Karras, and S.
755 Lindquist. 2012. Quantitative analysis of HSP90-client interactions reveals principles of
756 substrate recognition. Cell 150:987-1001.
757 Verba, K. A., R. Y. Wang, A. Arakawa, Y. Liu, M. Shirouzu, S. Yokoyama, and D. A.
758 Agard. 2016. Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes
759 an unfolded kinase. Science 352:1542-1547.
760 Xu, W., K. Beebe, J. D. Chavez, M. Boysen, Y. Lu, A. D. Zuehlke, D. Keramisanou, J. B.
761 Trepel, C. Prodromou, M. P. Mayer, J. E. Bruce, I. Gelis, and L. Neckers. 2019. Hsp90
762 middle domain phosphorylation initiates a complex conformational program to recruit the
763 ATPase-stimulating cochaperone Aha1. Nature communications 10:2574.

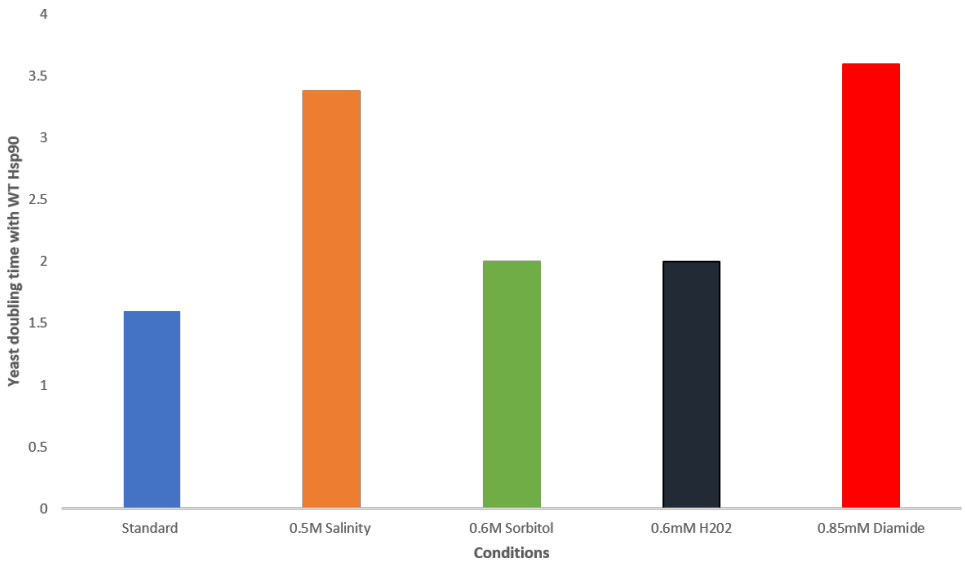
764 Xu, Y., M. A. Singer, and S. Lindquist. 1999. Maturation of the tyrosine kinase c-src as a
765 kinase and as a substrate depends on the molecular chaperone Hsp90. Proc. Natl. Acad. Sci.
766 U. S. A. 96:109-114.
767 Zabinsky, R. A., G. A. Mason, C. Queitsch, and D. F. Jarosz. 2019. It's not magic - Hsp90
768 and its effects on genetic and epigenetic variation. Semin. Cell Dev. Biol. 88:21-35.
769 Zhang, R., D. Luo, R. Miao, L. Bai, Q. Ge, W. C. Sessa, and W. Min. 2005. Hsp90–Akt
770 phosphorylates ASK1 and inhibits ASK1-mediated apoptosis. Oncogene 24:3954-3963.
771 Zhao, R., M. Davey, Y. C. Hsu, P. Kaplanek, A. Tong, A. B. Parsons, N. Krogan, G. Cagney,
772 D. Mai, J. Greenblatt, C. Boone, A. Emili, and W. A. Houry. 2005. Navigating the chaperone
773 network: an integrative map of physical and genetic interactions mediated by the hsp90
774 chaperone. Cell 120:715-727.

775

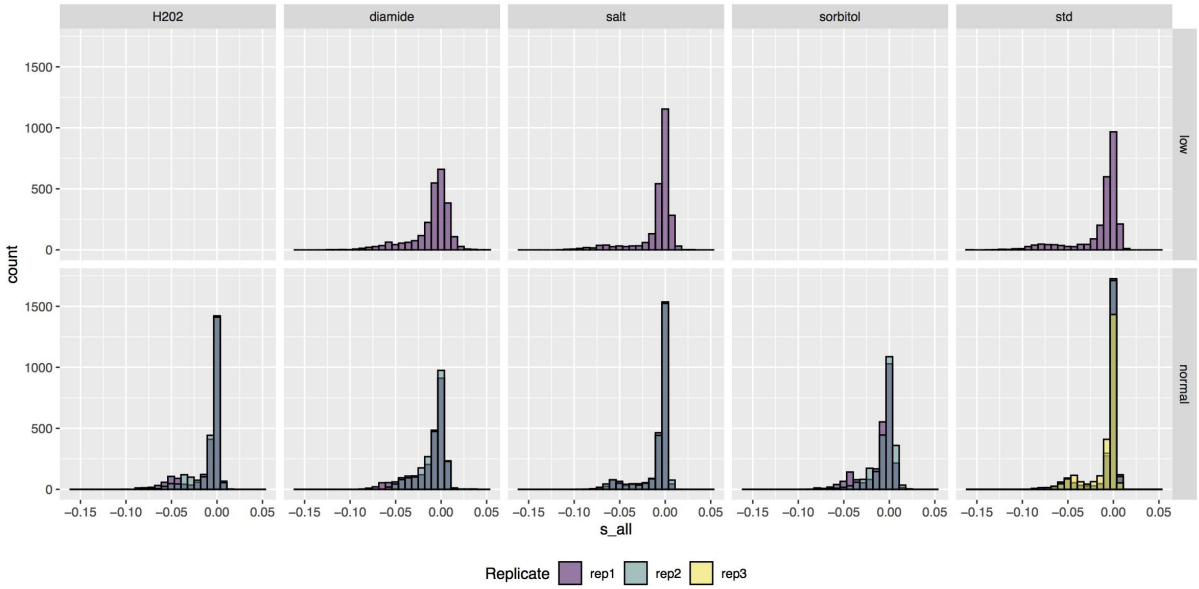
776

Supplementary Material

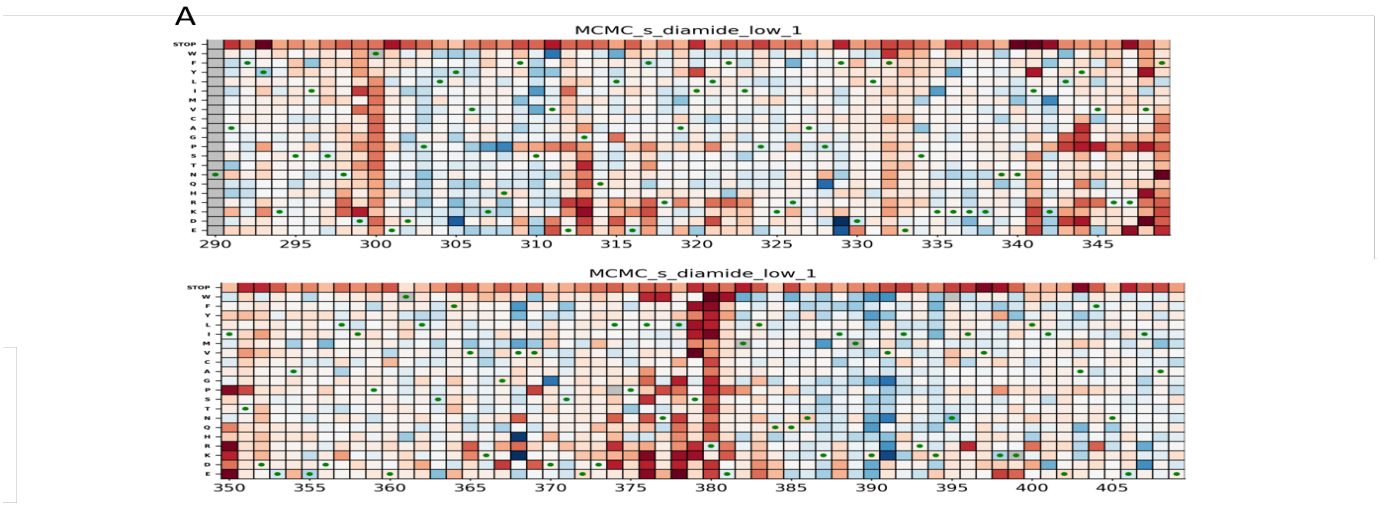
Supplementary Figures and Tables



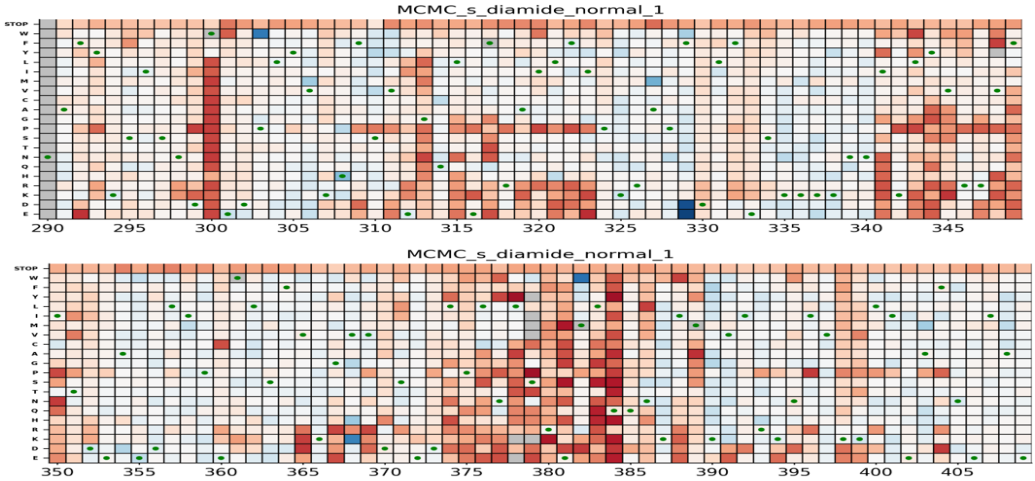
Supplementary figure 1. Doubling times for WT Hsp90 yeast growth under all environments. In general there is an increase in the doubling time in all environments tested compared to the standard environment. This increase is larger in high salinity and diamide environments, suggesting that these environments should present the highest stress load to the protein.



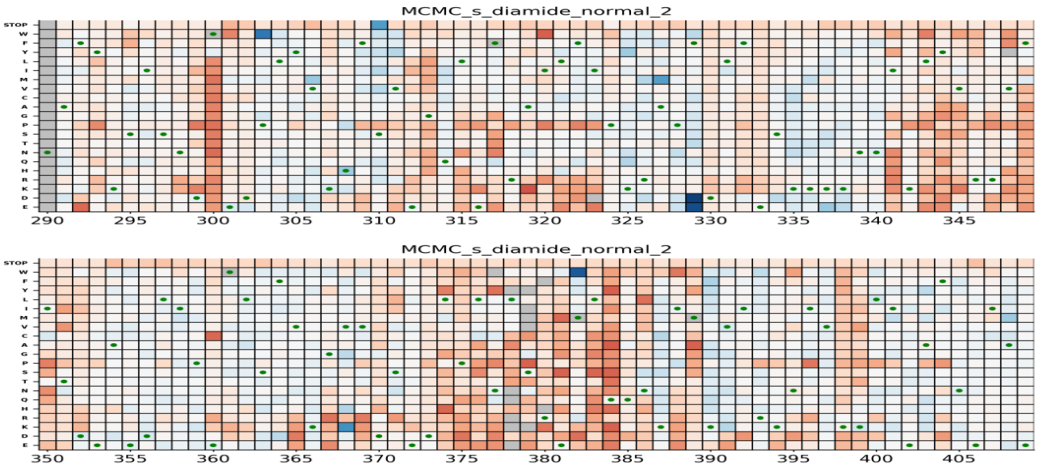
Supplementary Figure 2 Distribution of Fitness Effects (DFE) across environments and expression levels. Replicates within an environment largely overlap. There is a marked difference of the DFE in diamide compared to the other environments, especially at low expression levels. Interestingly, DFEs in H_20_2 and salt show a similar shape to the DFE in the standard environment, with a large number of mutations in the wild-type like region and a smaller peak of deleterious mutations, at which stop codons are located. In diamide and sorbitol, the DFE shows a lower number of wild-type like mutants, a larger variance, and a higher proportion of weakly and strongly deleterious mutations.



803

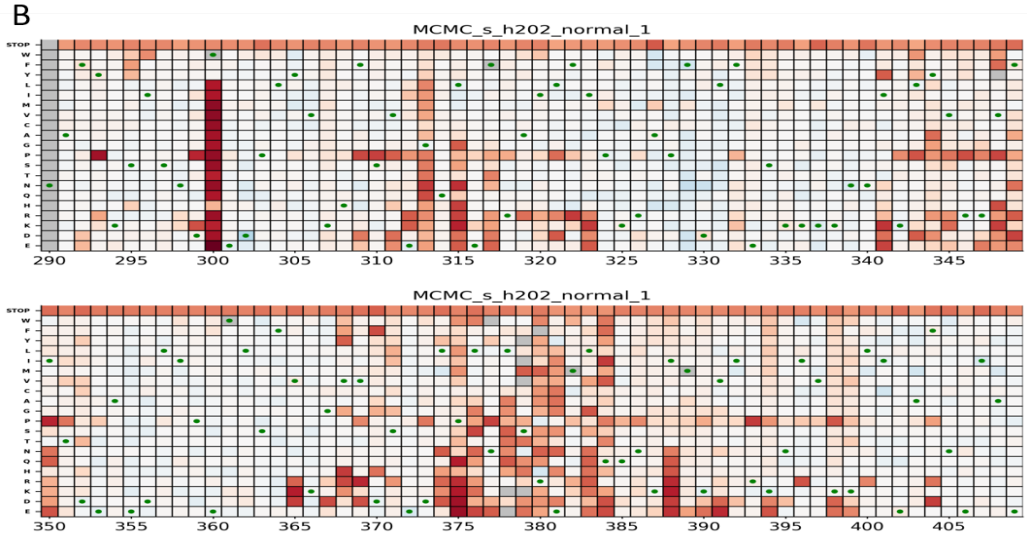

MCMC_s_diamide_normal_1
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_diamide_normal_1
350
355
360
365
370
375
380
385
390
395
400
405

804

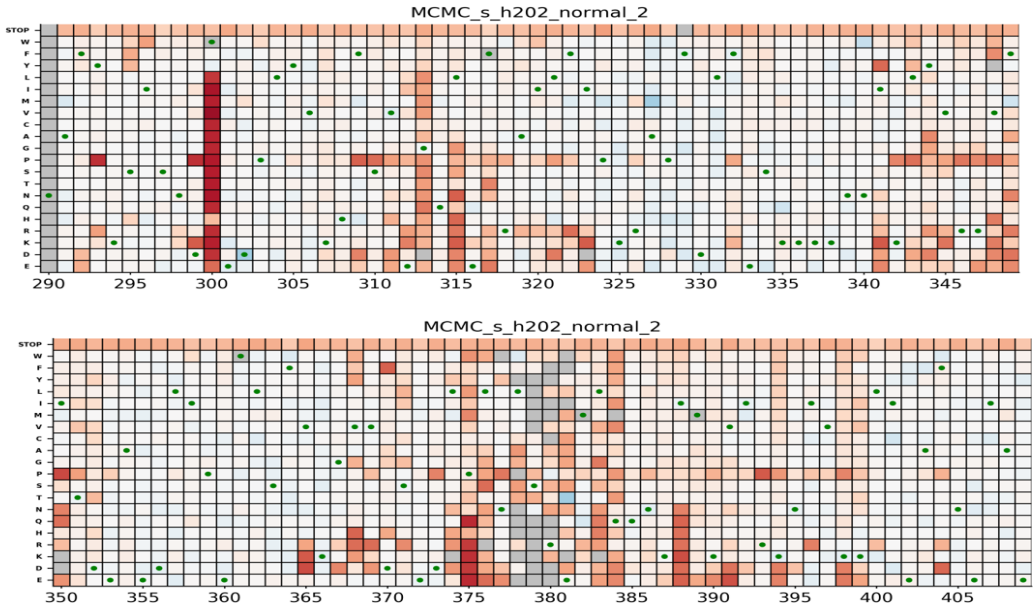

MCMC_s_diamide_normal_2
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_diamide_normal_2
350
355
360
365
370
375
380
385
390
395
400
405

805

B

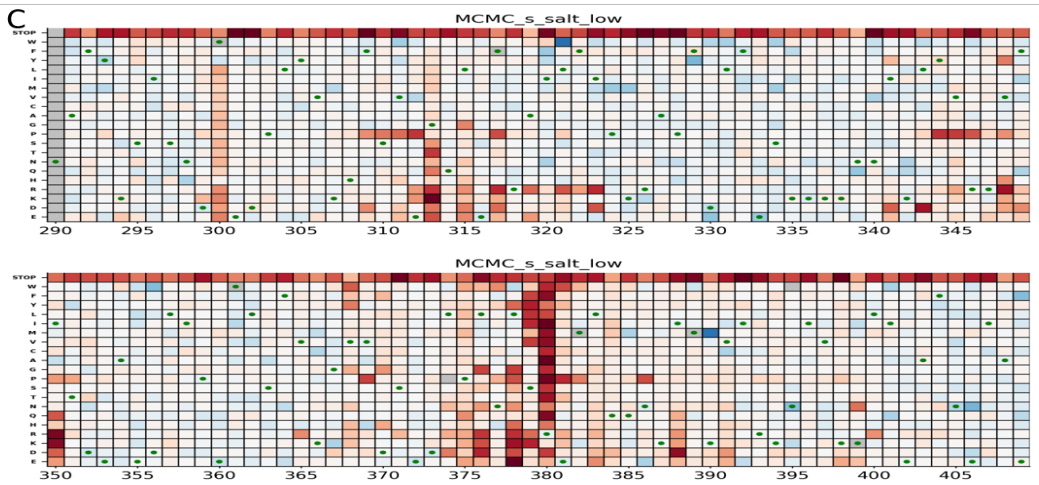

MCMC_s_h202_normal_1
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_h202_normal_1
350
355
360
365
370
375
380
385
390
395
400
405

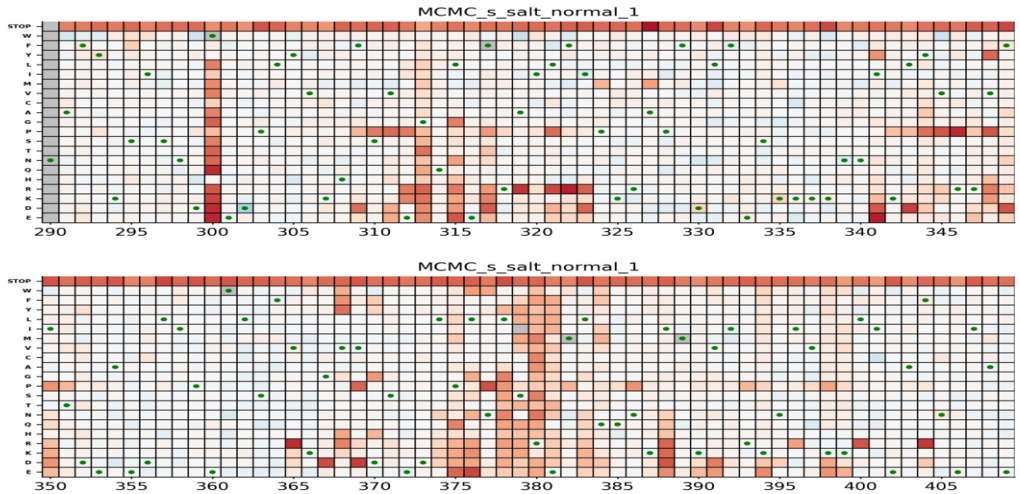
806

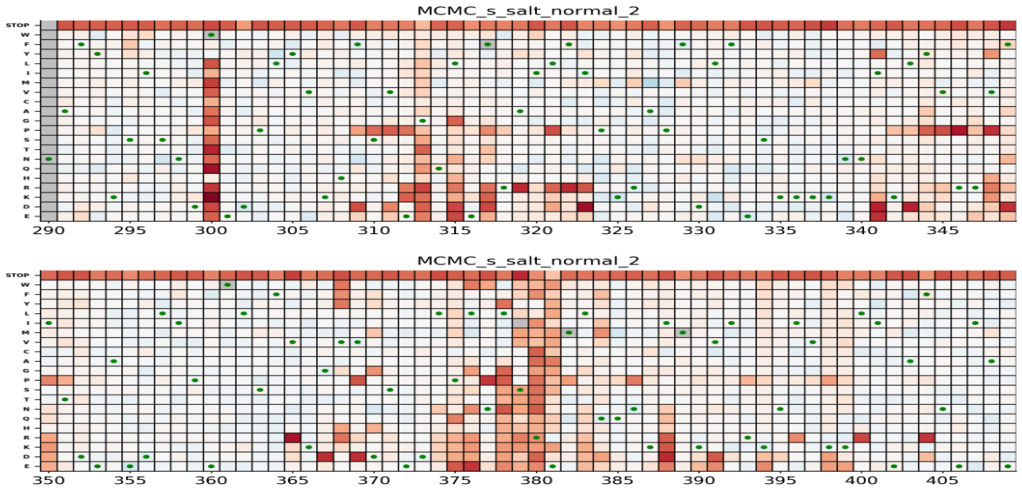

MCMC_s_h202_normal_2
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_h202_normal_2
350
355
360
365
370
375
380
385
390
395
400
405

807

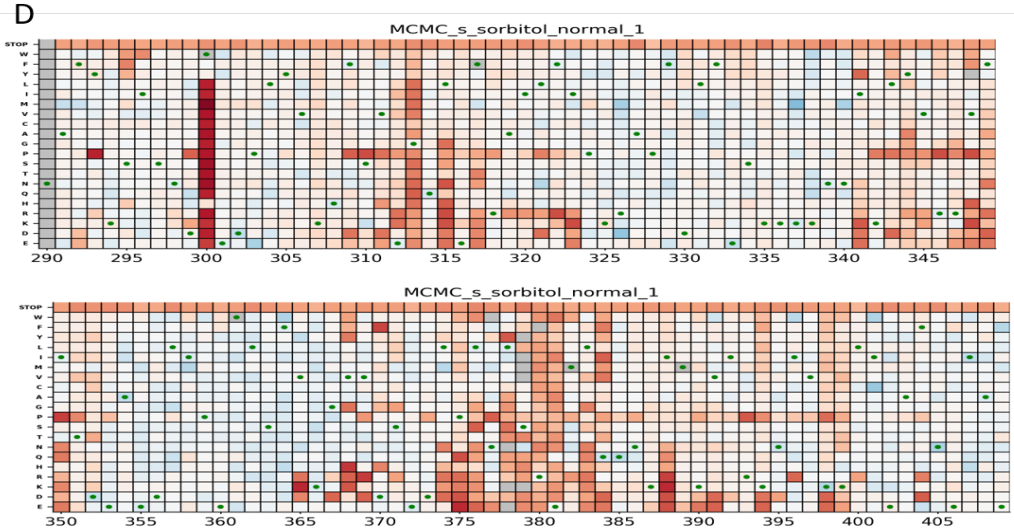
C


MCMC_s_salt_low
STOP
290 295 300 305 310 315 320 325 330 335 340 345
MCMC_s_salt_low
STOP
350 355 360 365 370 375 380 385 390 395 400 405

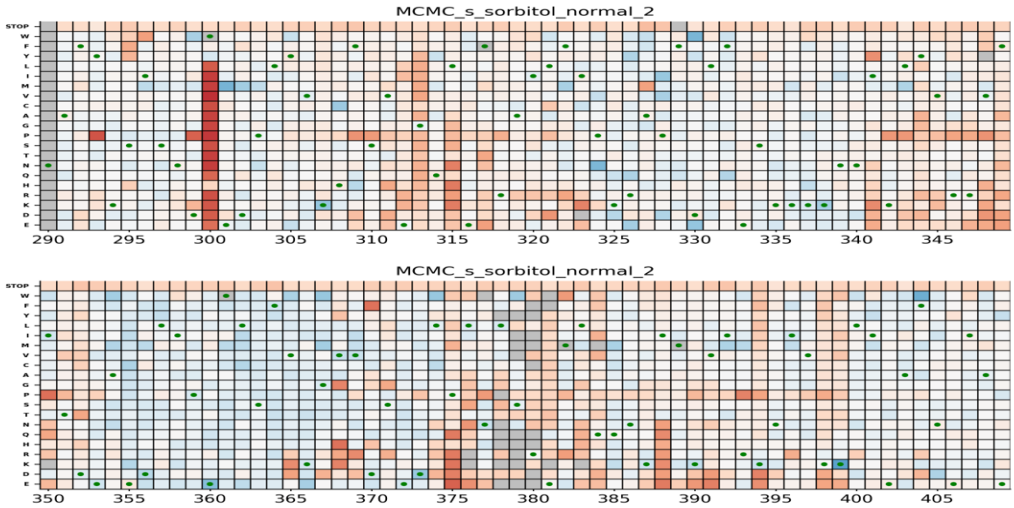

MCMC_s_salt_normal_1
STOP
290 295 300 305 310 315 320 325 330 335 340 345
MCMC_s_salt_normal_1
STOP
350 355 360 365 370 375 380 385 390 395 400 405


MCMC_s_salt_normal_2
STOP
290 295 300 305 310 315 320 325 330 335 340 345
MCMC_s_salt_normal_2
STOP
350 355 360 365 370 375 380 385 390 395 400 405

D

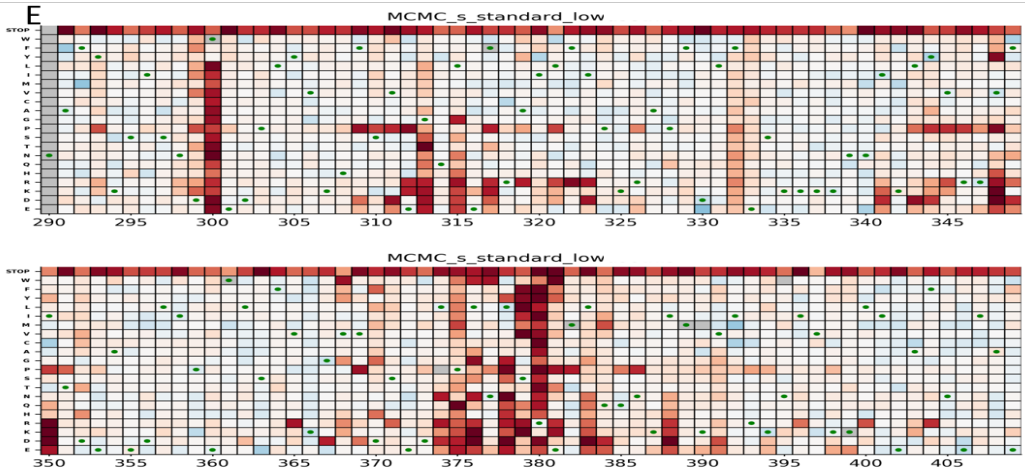

MCMC_s_sorbitol_normal_1
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_sorbitol_normal_1
350
355
360
365
370
375
380
385
390
395
400
405

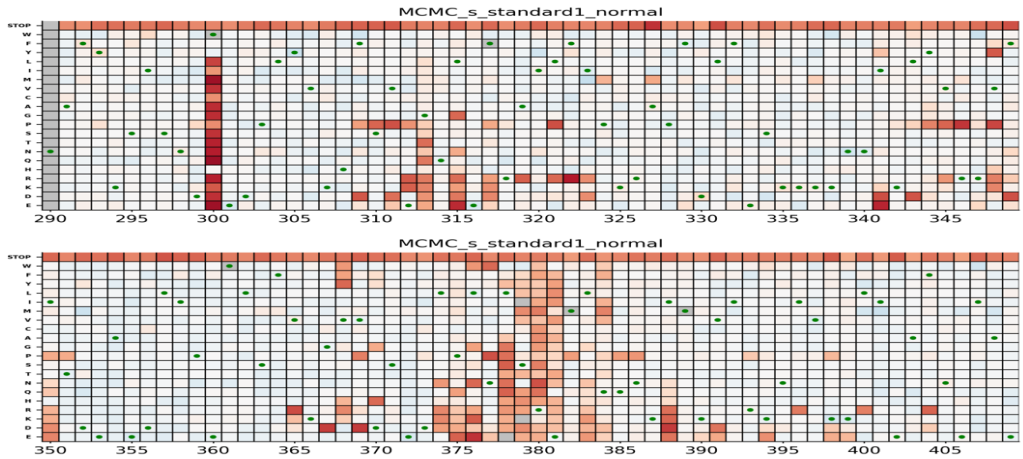
811

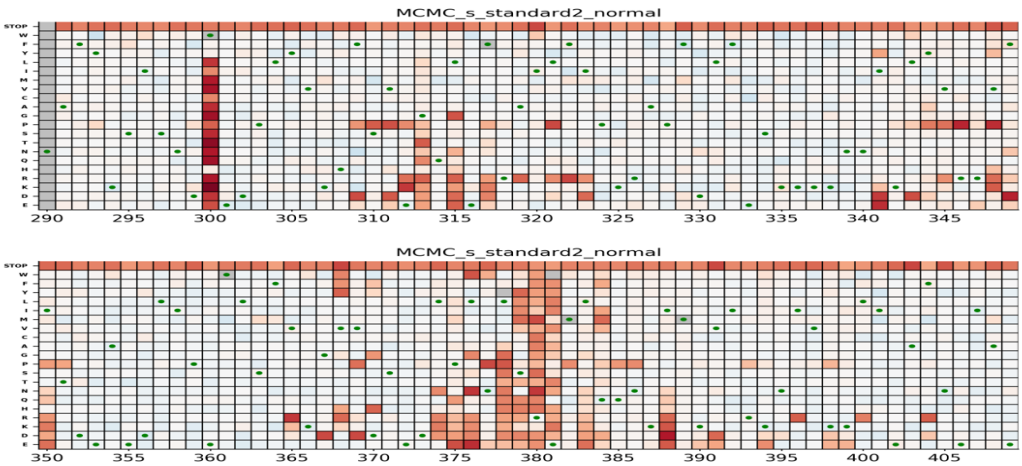

MCMC_s_sorbitol_normal_2
STOP
W
F
Y
L
I
M
V
C
A
G
P
S
T
N
Q
H
R
K
D
E
290
295
300
305
310
315
320
325
330
335
340
345
MCMC_s_sorbitol_normal_2
350
355
360
365
370
375
380
385
390
395
400
405

812

E



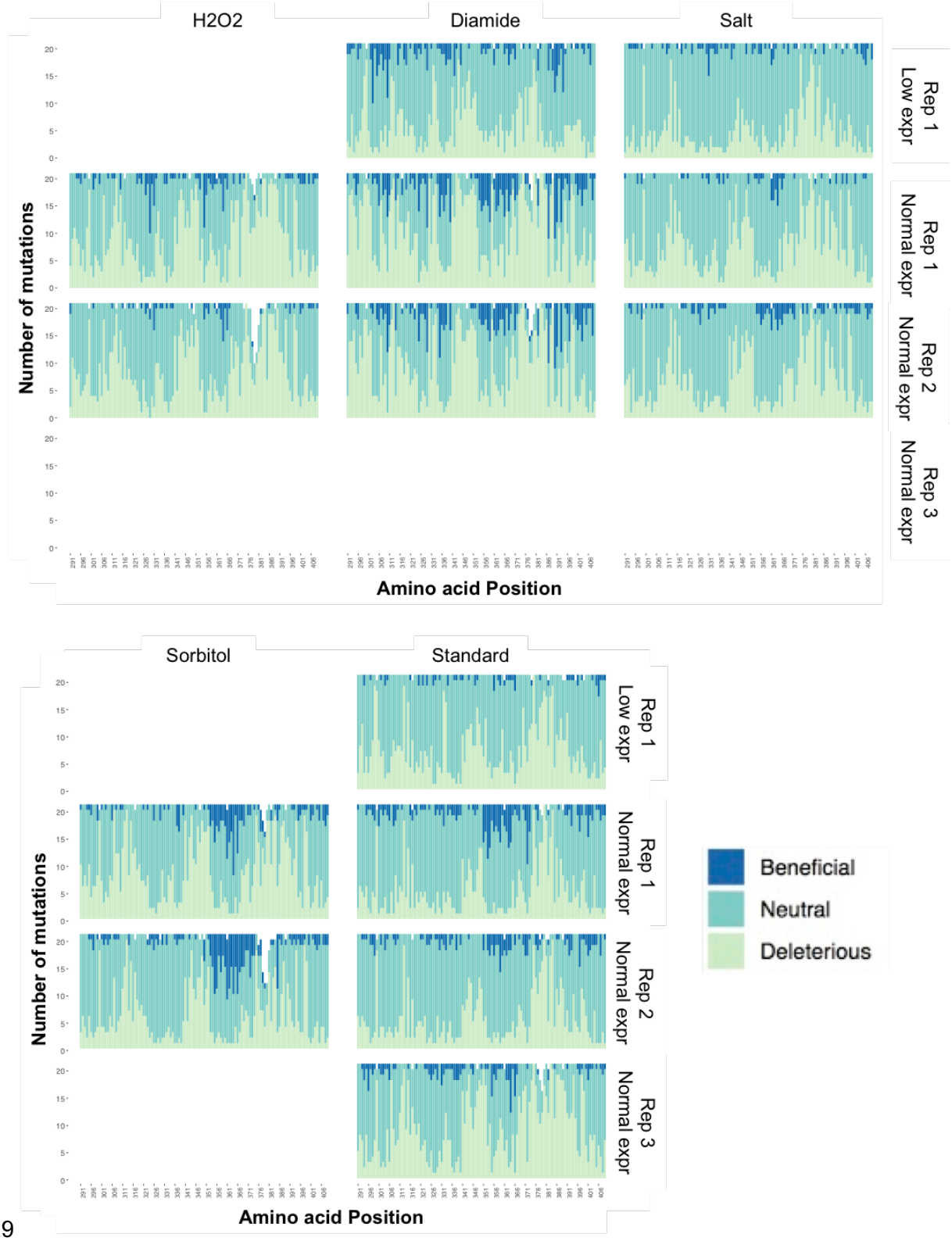




= No observation

= Wildtype amino acid

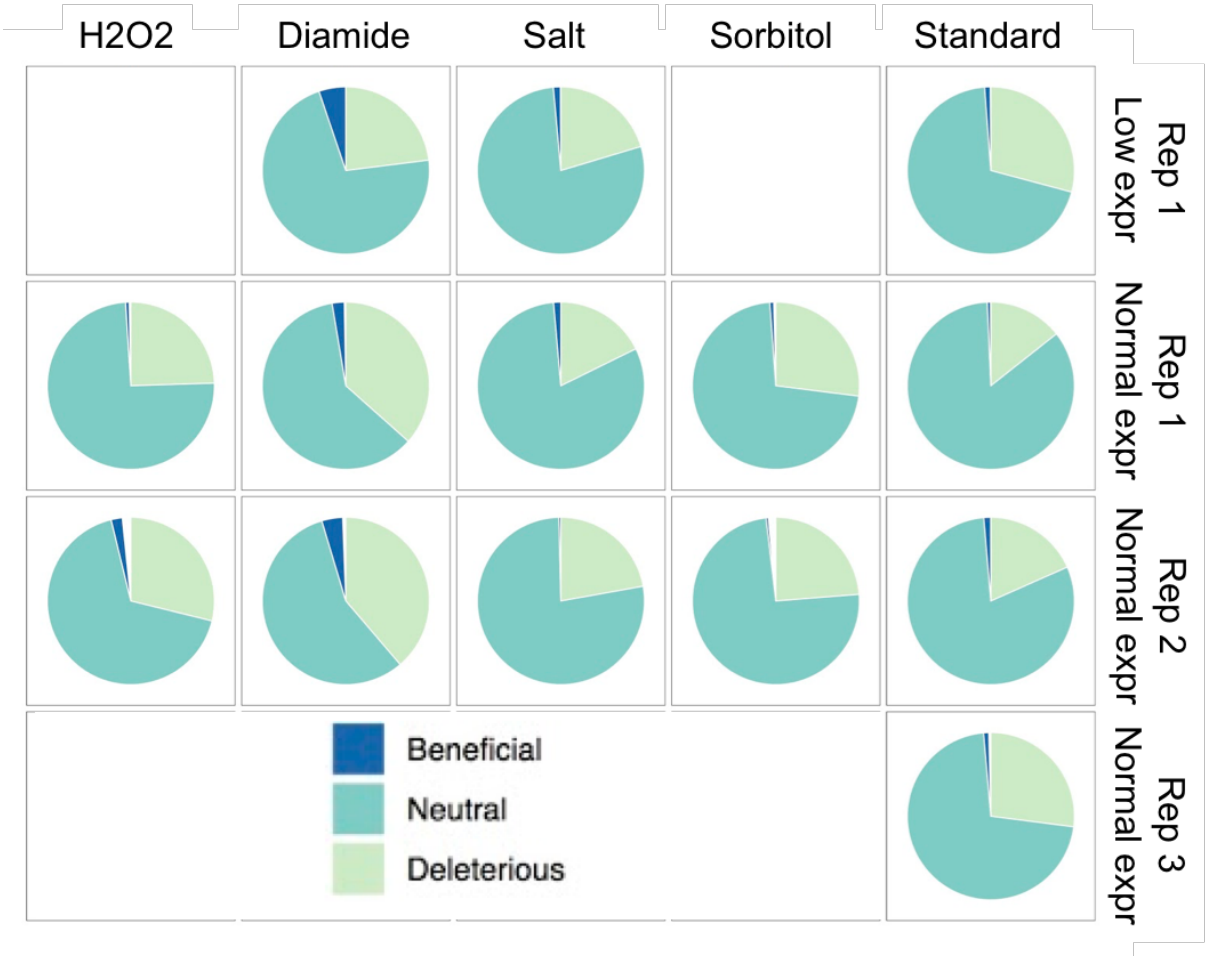
816 **Supplementary Figure 3**. **Heatmaps of mutant fitness effects show that the majority of**
817 **mutations are wild type under all conditions, with the exception of amino acid regions**
818 **~300, 315-320, 347 and 375-390 that show mostly deleterious mutant fitness effects.**
819 **Beneficial mutations were found under all conditions but in different regions of the**
820 **middle domain.** (A) Diamide heatmaps under low and normal expression. (B) H_20_2
821 heatmaps under normal expression. (C) Salt heatmaps under low and normal expression. (D)
822 Sorbitol heatmaps under normal expression. (E) Standard heatmaps under low and normal
823 expression.
824
825
826
827
828



Supplementary Figure 4

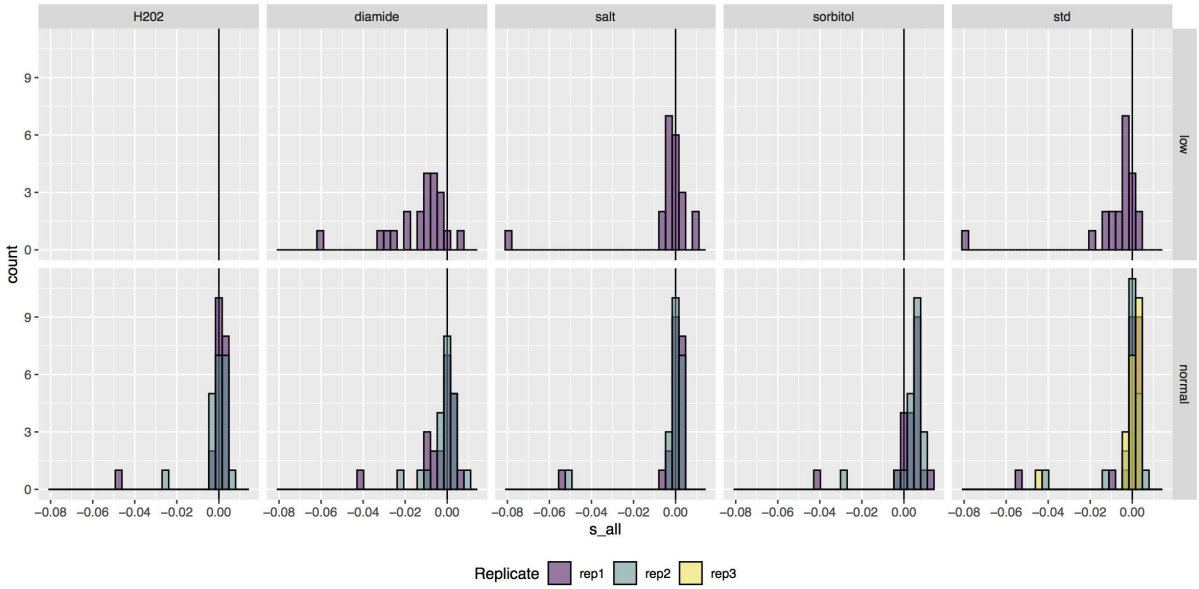
Proportions of deleterious, beneficial and wild-type like mutations along the middle domain of Hsp90. The x-axis corresponds to the amino acid positions studied and the y axis

to the proportion of deleterious, beneficial and neutral mutations at each amino acid position. The rows indicate expression level and replicate, the columns the different environments.



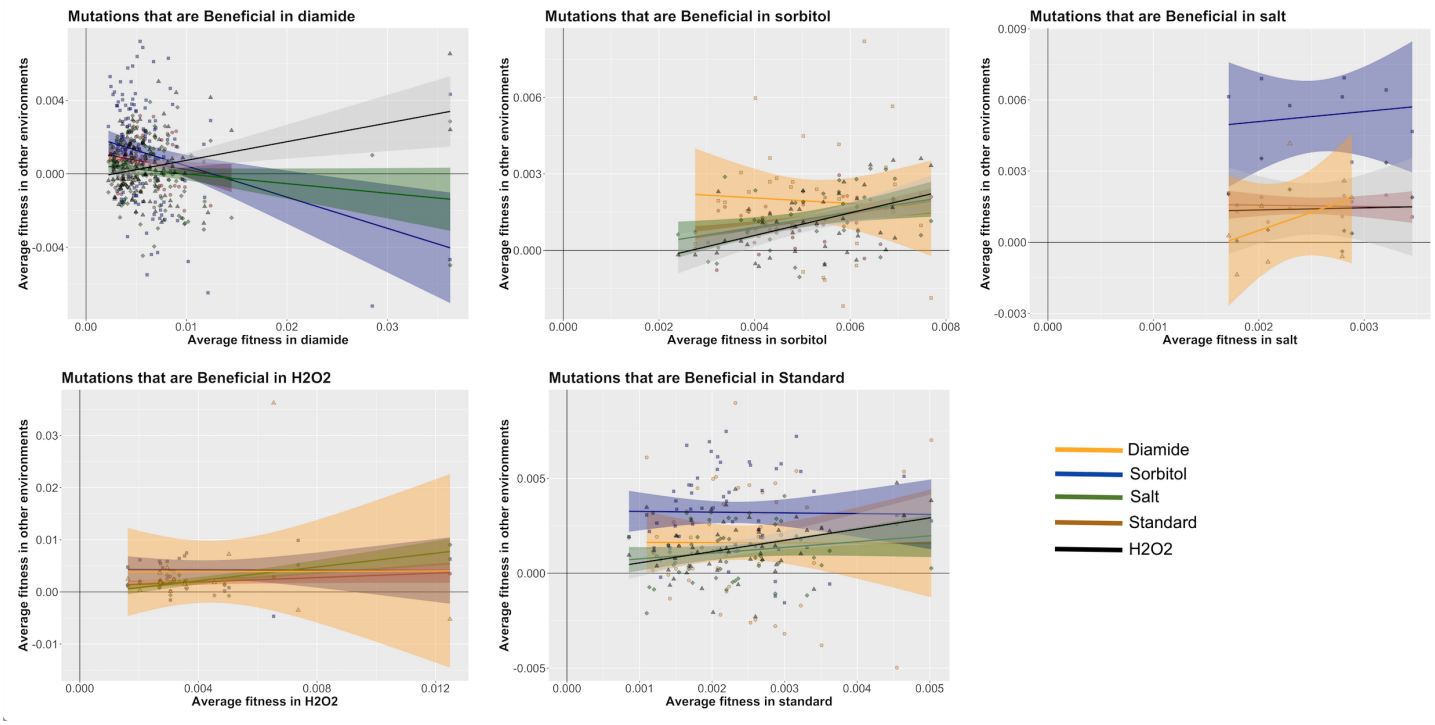
Supplementary Figure 5

Summary of the proportions of beneficial, deleterious and neutral mutations in each environment/replicate/expression level. In general we see few beneficial mutations, except in diamide, especially at low expression. In diamide, there is also a larger proportion of deleterious mutations in comparison with other environments.



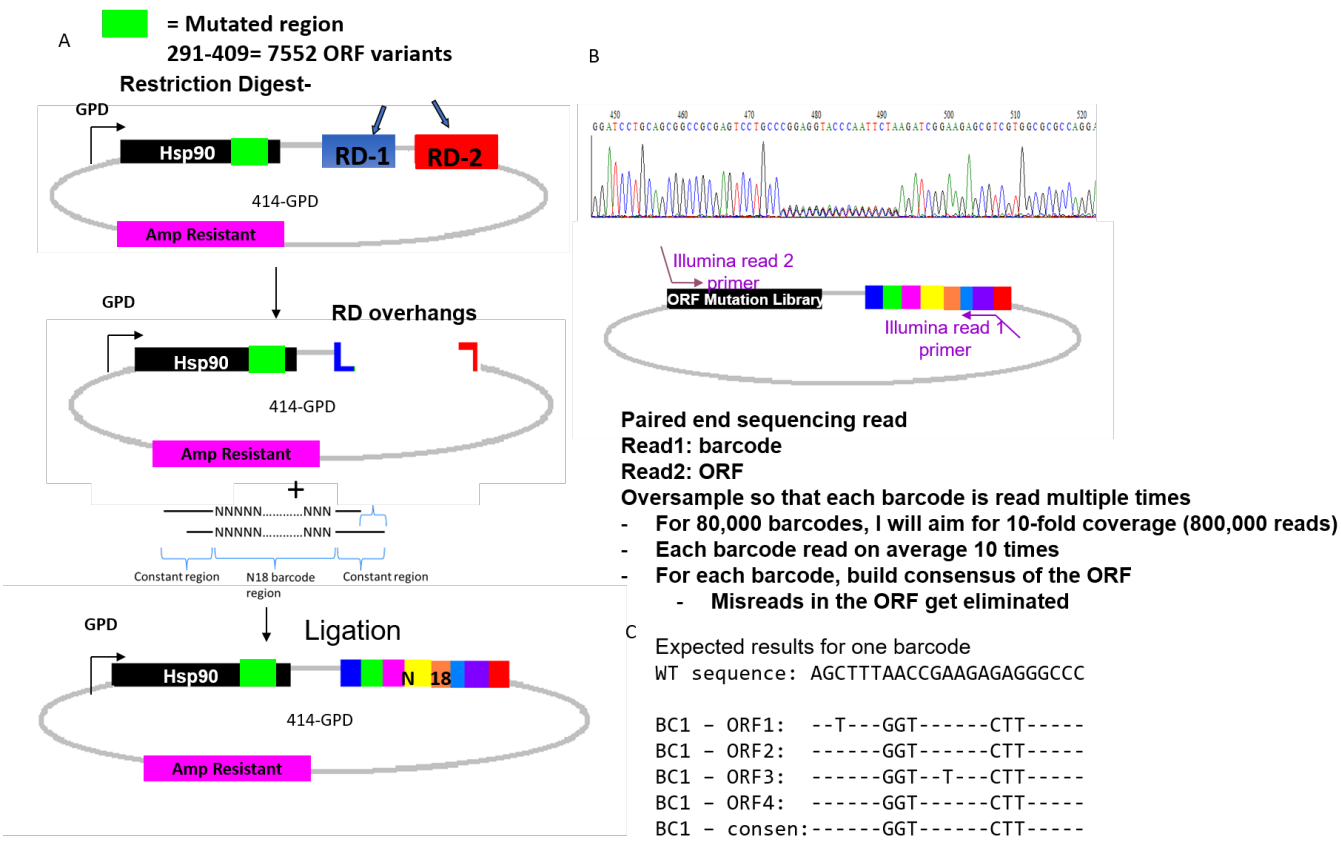
Supplementary Figure 6

Distribution of fitness effects at position 364. Compared with the DFE of the whole middle domain (Supplementary Figure 1), there is a clear enrichment of beneficial mutations in position 364, especially at normal expression levels.



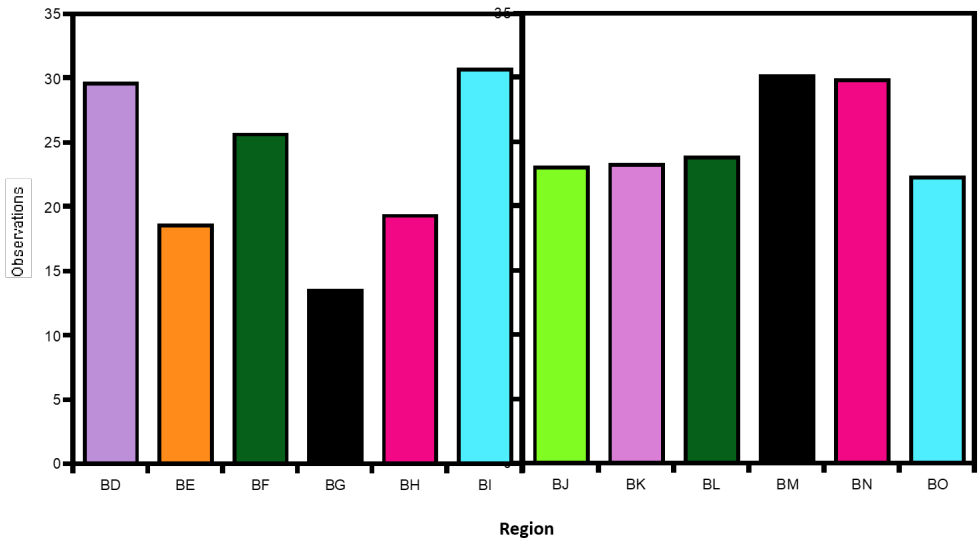
Supplementary Figure 7

854 **Cost of adaptation in different environments.** Each plot represents the average fitness
855 effect of mutations that are beneficial in a specific environment (x-axis) and the average
856 fitness effect in other environment (y-axis). Focal environments are indicated by the color of
857 the lines and points. Note that the x and y scales vary between panels. To improve
858 visualization, we calculated the average fitness effects between the two (or three replicates)
859 and mutations were recategorized based on whether mutations were beneficial in two
860 replicates (beneficial mutations), beneficial in one replicate and neutral in another
861 (beneficial-neutral), neutral in both replicates (neutral mutations), neutral in one replicate and
862 deleterious in the other (neutral-deleterious mutations), deleterious in both replicates
863 (deleterious) or deleterious in one replicate in beneficial in the other (beneficial-deleterious
864 mutations).
865

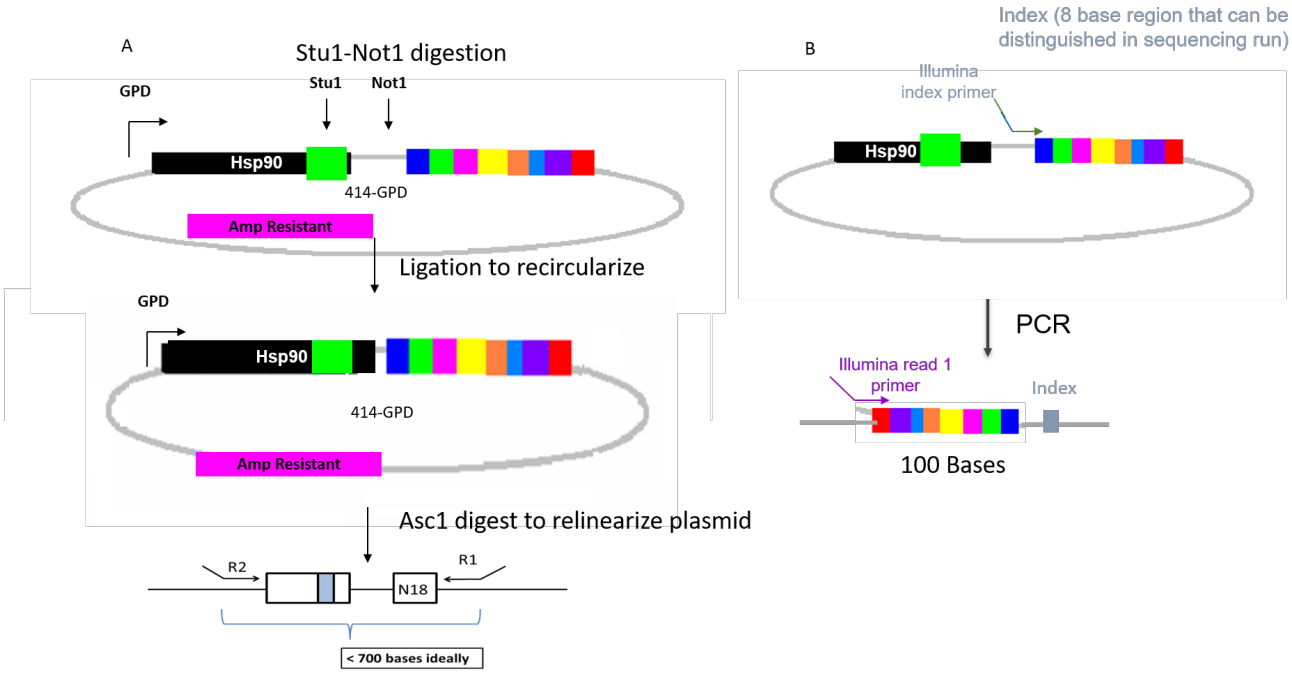


866

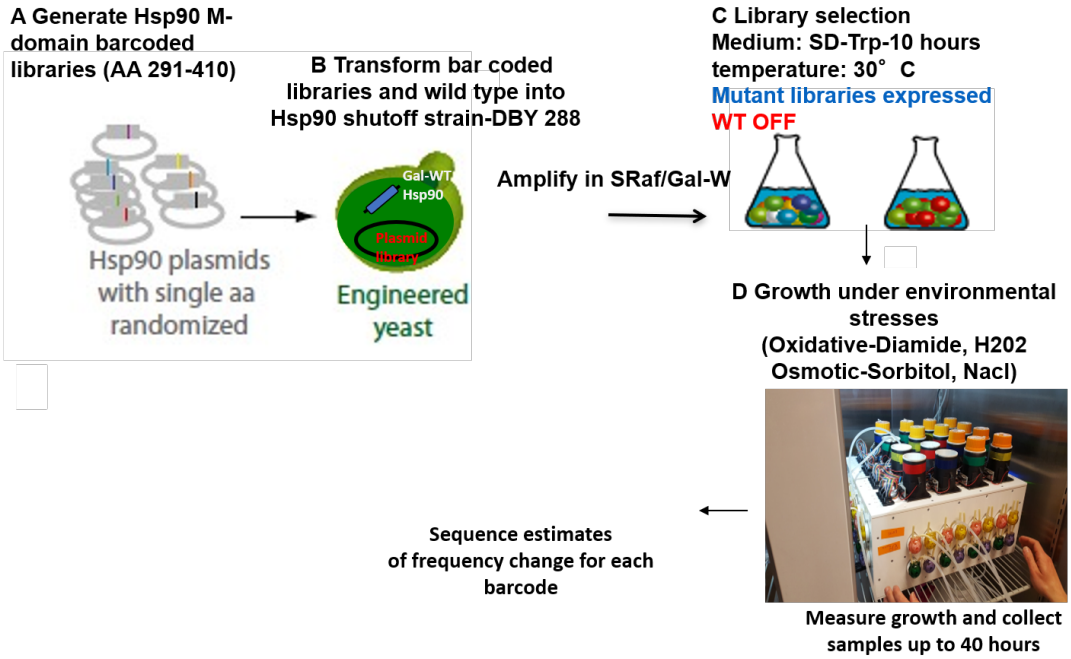
867 **Supplementary figure 8**. Barcoding mutant library strategy. (A) Schematic of how to
868 incorporate barcodes into plasmid libraries using restriction digests and ligation. (B) Sanger
869 sequence TRACE result of barcoded region of the library at positions 475-493. (C)
870 Schematic of how to map barcodes to ORF mutants.



Supplementary figure 9.Average number of barcodes per mutant for each 9 amino acid window shows an average of 10 barcodes associated with each variant .



Supplementary figure 10. Strategy to prepare barcoded mutant libraries for sequencing. (A) Schematic of how to prepare barcoded mutant libraries for paired end deep sequencing.(B) Schematic of how to include an index primer to distinguish each time point and final sequencing read.



Supplementary figure 11: Systematic approach to measure the adaptive potential of the middle domain of yeast Hsp90 (amino acids 291-409).

Environment	Replicate	Expression	% Beneficial	% Neutral	% Deleterious	Total Mutations
H202	1	normal	5.03 (125)	50.2 (1248)	44.77 (1113)	2486
H202	2	normal	2.65 (65)	54.5 (1337)	42.85 (1051)	2453
Diamide	1	low	7.9 (197)	63.54 (1584)	28.56 (712)	2493
Diamide	1	normal	12.34 (307)	34.74 (864)	52.92 (1316)	2487
Diamide	2	normal	10.73 (266)	41.19 (1021)	48.08 (1192)	2479
High Salinity	1	low	3.53 (88)	72.59 (1809)	23.88 (595)	2492
High Salinity	1	normal	2.57 (64)	59.58 (1486)	37.85 (944)	2494
High Salinity	2	normal	4.85 (121)	59.34 (1480)	35.81 (893)	2494
Sorbitol	1	normal	7.09 (176)	51.37 (1276)	41.55 (1032)	2484
Sorbitol	2	normal	9.97 (245)	60.09 (1477)	29.94 (736)	2458
Standard	1	low	2.53 (63)	62.55 (1558)	34.93 (870)	2491
Standard	1	normal	9.03 (225)	65.53 (1633)	25.44 (634)	2492
Standard	2	normal	6.14 (153)	66.75 (1664)	27.12 (676)	2493
Standard	3	normal	5.19 (129)	49.6 (1232)	45.21 (1123)	2484

Supplementary Table 1

Number of beneficial, neutral and deleterious mutations across environments, expression levels and replicates. Gray cells corresponds to low expression level mutants (from Flynn et al 2019). We observed a large proportion of wild-type like mutations in all environments.

			Low Expression			Normal Expression										
						Diamide		Salt		Standard			H2O2		Sorbitol	
			Standard	Diamide	Salt	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep3	Rep1	Rep2	Rep1	Rep2
Low Expression	Standard															
	Diamide		0.77													
	Salt		0.88	0.70												
Normal Expression	Diamide	Rep1	0.61	0.61	0.51											
		Rep2	0.48	0.50	0.36	0.96										
	Salt	Rep1	0.86	0.70	0.85	0.64	0.48									
		Rep2	0.87	0.71	0.85	0.65	0.50	0.97								
	Standard	Rep1	0.86	0.69	0.83	0.63	0.47	0.96	0.95							
		Rep2	0.87	0.70	0.83	0.64	0.48	0.94	0.96	0.97						
		Rep3	0.78	0.63	0.68	0.81	0.71	0.83	0.83	0.82	0.83					
	H2O2	Rep1	0.80	0.64	0.72	0.79	0.68	0.86	0.86	0.86	0.86	0.99				
		Rep2	0.76	0.59	0.65	0.78	0.70	0.78	0.80	0.77	0.79	0.97	0.96			
	Sorbitol	Rep1	0.79	0.64	0.70	0.76	0.66	0.84	0.85	0.82	0.84	0.95	0.95	0.94		
		Rep2	0.69	0.54	0.58	0.69	0.62	0.73	0.73	0.71	0.72	0.90	0.88	0.90	0.87	

Cor between replicates in normal expression
Cor between expression levels (same environment)

Supplementary Table 2

Pearson correlation of fitness effects across replicates, environments, and expression levels. Blue indicates weaker correlations and red stronger correlations. Full black lines indicate correlations between expression levels and double lines indicate correlations between replicates of the same environment. In general the lowest correlations are seen for diamide in high and low expression and salt in low expression. Correlations between replicates are in general strong, except for standard replicate 3. Most correlations are strong, which suggests that in general the fitness effects of mutations should not undergo large changes between environments.

				Deleterious in														
				Low Expression			Normal Expression											
							Diamide		Salt		Standard			H2O2		Sorbitol	
				Standard	Diamide	Salt	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2	Rep3	Rep1	Rep2	Rep1	Rep2
Beneficials in	Low Expression	Standard		63	12.7 (8)	6.35 (4)	14.29 (9)	7.94 (5)	14.29 (9)	6.35 (4)	6.35 (4)	6.35 (4)	9.52 (6)	7.94 (5)	6.35 (4)	4.76 (3)	4.76 (3)
		Diamide		32.99 (65)	197	23.86 (47)	28.43 (56)	23.35 (46)	30.46 (60)	31.47 (62)	16.75 (33)	21.83 (43)	39.59 (78)	45.18 (89)	44.16 (87)	37.56 (74)	28.93 (57)
		Salt		15.91 (14)	27.27 (24)	88	37.5 (33)	27.27 (24)	25 (22)	13.64 (12)	7.95 (7)	6.82 (6)	23.86 (21)	21.59 (19)	23.86 (21)	17.05 (15)	11.36 (10)
	Normal Expression	Diamide	Rep1	19.87 (61)	9.12 (28)	11.4 (35)	307	0 (0)	16.61 (51)	18.24 (56)	6.84 (21)	8.14 (25)	24.43 (75)	26.38 (81)	28.01 (86)	22.8 (70)	14.01 (43)
			Rep2	21.05 (56)	10.9 (29)	12.41 (33)	0.38 (1)	266	19.17 (51)	18.05 (48)	9.02 (24)	9.02 (24)	28.95 (77)	30.83 (82)	30.45 (81)	25.19 (67)	15.79 (42)
		Salt	Rep1	10.94 (7)	12.5 (8)	0 (0)	26.56 (17)	25 (16)	64	6.25 (4)	0 (0)	4.69 (3)	1.56 (1)	3.12 (2)	9.38 (6)	7.81 (5)	1.56 (1)
			Rep2	16.53 (20)	15.7 (19)	4.13 (5)	28.93 (35)	23.14 (28)	6.61 (8)	121	4.13 (5)	0.83 (1)	9.09 (11)	11.57 (14)	11.57 (14)	6.61 (8)	0.83 (1)
		Standard	Rep1	17.33 (39)	18.67 (42)	9.78 (22)	28.89 (65)	27.11 (61)	7.56 (17)	10.22 (23)	225	2.22 (5)	20.44 (46)	20 (45)	20.44 (46)	19.11 (43)	8.89 (20)
			Rep2	21.57 (33)	19.61 (30)	8.5 (13)	33.99 (52)	32.68 (50)	9.15 (14)	5.88 (9)	3.92 (6)	153	24.18 (37)	24.18 (37)	19.61 (30)	12.42 (19)	5.88 (9)
			Rep3	12.4 (16)	13.95 (18)	2.33 (3)	22.48 (29)	19.38 (25)	4.65 (6)	4.65 (6)	2.33 (3)	3.88 (5)	129	0 (0)	10.08 (13)	2.33 (3)	0 (0)
		H2O2	Rep1	11.2 (14)	17.6 (22)	5.6 (7)	24 (30)	17.6 (22)	4.8 (6)	3.2 (4)	2.4 (3)	4 (5)	0.8 (1)	125	4.8 (6)	6.4 (8)	1.6 (2)
			Rep2	12.31 (8)	10.77 (7)	1.54 (1)	21.54 (14)	16.92 (11)	4.62 (3)	1.54 (1)	6.15 (4)	1.54 (1)	4.62 (3)	0 (0)	65	3.08 (2)	1.54 (1)
		Sorbitol	Rep1	15.91 (28)	15.34 (27)	5.11 (9)	27.84 (49)	23.3 (41)	5.68 (10)	2.27 (4)	2.27 (4)	2.84 (5)	5.11 (9)	7.39 (13)	10.23 (18)	176	3.98 (7)
			Rep2	15.1 (37)	16.33 (40)	6.94 (17)	34.29 (84)	29.8 (73)	7.35 (18)	3.67 (9)	4.9 (12)	4.08 (10)	12.24 (30)	16.73 (41)	13.06 (32)	15.1 (37)	245

Percentage of beneficials that have deleterious effect between replicates
Percentage of beneficials that have deleterious effect between levels of expression
Number of beneficials in each environment

Supplementary Table 3

Percentage (and number) of beneficials and their cost across environments. Beneficials in a focal environment (rows) are deleterious in the environment indicated by the column. The diagonal of the matrix indicates the number of beneficials found in each environment/replicate. Yellow shading indicates "spurious costs of adaptation" between pairs

913 of replicates, which is a proxy for the expected classification error. In general, this error is
914 low (between 0.38% and 6.61%), except for Replicate 3 in the standard environment and
915 replicate 2 from Sorbitol, where many beneficial mutations in one replicate are classified as
916 deleterious in the other. Green highlights the comparisons between levels of expression for
917 standard, diamide and salt environments. Interestingly, for diamide and salt, the percentage of
918 beneficials at normal levels of expression that are deleterious at low expression is low, but
919 there is a larger proportion of beneficials at low expression that is deleterious at high
920 expression. This pattern in inverted for the standard environment, which is consistent with the
921 results from Jiang et al 2013.
922