

On the (un)predictability of a large intragenic fitness landscape

Claudia Bank^{a,b,c,1}, Sebastian Matuszewski^{b,c,1}, Ryan T. Hietpas^{d,e}, and Jeffrey D. Jensen^{b,c,2,3}

^aInstituto Gulbenkian de Ciência, 2780-156 Oeiras, Portugal; ^bSchool of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; ^cSwiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; ^dEli Lilly and Company, Indianapolis, IN 46225; and ^eDepartment of Biochemistry & Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605

Edited by Andrew G. Clark, Cornell University, Ithaca, NY, and approved October 11, 2016 (received for review August 2, 2016)

The study of fitness landscapes, which aims at mapping genotypes to fitness, is receiving ever-increasing attention. Novel experimental approaches combined with next-generation sequencing (NGS) methods enable accurate and extensive studies of the fitness effects of mutations, allowing us to test theoretical predictions and improve our understanding of the shape of the true underlying fitness landscape and its implications for the predictability and repeatability of evolution. Here, we present a uniquely large multiallelic fitness landscape comprising 640 engineered mutants that represent all possible combinations of 13 amino acid-changing mutations at 6 sites in the heat-shock protein Hsp90 in *Saccharomyces cerevisiae* under elevated salinity. Despite a prevalent pattern of negative epistasis in the landscape, we find that the global fitness peak is reached via four positively epistatic mutations. Combining traditional and extending recently proposed theoretical and statistical approaches, we quantify features of the global multiallelic fitness landscape. Using subsets of the data, we demonstrate that extrapolation beyond a known part of the landscape is difficult owing to both local ruggedness and amino acid-specific epistatic hotspots and that inference is additionally confounded by the nonrandom choice of mutations for experimental fitness landscapes.

evolution | adaptation | epistasis | fitness landscape | mutagenesis

Since first proposed by Sewall Wright in 1932 (1), the idea of a fitness landscape relating genotype (or phenotype) to the reproductive success of an individual has inspired evolutionary biologists and mathematicians alike (2–4). With the advancement of molecular and systems biology toward large and accurate datasets, the fitness landscape concept has received increasing attention across other subfields of biology (5–9). The shape of the fitness landscape carries information on the repeatability and predictability of evolution, the potential for adaptation, the importance of genetic drift, the likelihood of convergent and parallel evolution, and the degree of optimization that is (theoretically) achievable (4). Unfortunately, the dimensionality of a complete fitness landscape of an organism—that is, a mapping of all possible combinations of mutations to their respective fitness effects—is much too high to be assessed experimentally. With the development of experimental approaches that allow for the assessment of full fitness landscapes of tens to hundreds of mutations, there is growing interest in statistics that capture the features of the landscape and that relate an experimental landscape to theoretical landscapes of similar architecture, which have been studied extensively (10). It is, however, unclear whether this categorization allows for an extrapolation to unknown parts of the landscape, which would be the first step toward quantifying predictability—an advancement that would yield impacts far beyond the field of evolutionary biology, in particular for the clinical study of drug-resistance evolution in pathogens and the development of effective vaccine and treatment strategies (8).

Existing research in this rapidly growing field comes from two sides. Firstly, different empirical landscapes have been assessed (reviewed in ref. 4), generally based on the combination

of previously observed beneficial mutations or on the dissection of an observed adaptive walk (i.e., a combination of mutations that have been observed to be beneficial in concert). Secondly, theoretical research has proposed different landscape architectures [such as the House-of-Cards (HoC), the Kauffman NK (NK), and the Rough Mount Fuji (RMF) model], studied their respective properties, and developed a number of statistics that characterize the landscape and quantify the expected degree of epistasis (i.e., interaction effects between mutations) (10–14).

The picture that emerges from these studies is mixed, reporting both smooth (15) and rugged (16, 17) landscapes with both positive epistasis [i.e., two mutations in concert are more advantageous than expected; (18)] and negative epistasis [i.e., two mutations in concert are more deleterious than expected (refs. 19 and 20 but see refs. 21 and 22)]. Current statistical approaches have been used to rank the existing landscapes by certain features (10, 14) and to assess whether the landscapes are compatible with Fisher's Geometric Model (23). A crucial remaining question is the extent to which the nonrandom choice of mutations for the experiment affects the topography of the landscape and whether the local topography is indeed informative as to the rest of the landscape.

Here, we present an intragenic fitness landscape of 640 amino acid-changing mutations in the heat-shock protein Hsp90 in *Saccharomyces cerevisiae* in a challenging environment imposed by

Significance

The study of fitness landscapes is fundamentally concerned with understanding the relative roles of stochastic and deterministic processes in adaptive evolution. Here, the authors present a uniquely large and complete multiallelic intragenic fitness landscape of 640 systematically engineered mutations in the heat-shock protein Hsp90 in yeast. Using a combination of traditional and recently proposed theoretical approaches, they study the accessibility of the global fitness peak and the potential for predictability of the fitness landscape topography. They report local ruggedness of the landscape and the existence of epistatic hotspot mutations, which together make extrapolation and hence predictability inherently difficult if mutation-specific information is not considered.

Author contributions: C.B., R.T.H., and J.D.J. designed research; R.T.H. performed research; C.B. and S.M. analyzed data; and C.B., S.M., R.T.H., and J.D.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Dryad Digital Repository (dx.doi.org/10.5061/dryad.th0rj).

¹C.B. and S.M. contributed equally to this work.

²Present address: Center for Evolution & Medicine, School of Life Sciences, Arizona State University, Tempe, AZ 85281.

³To whom correspondence should be addressed. Email: jeffrey.d.jensen@asu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612676113/-DCSupplemental.

high salinity. With all possible combinations of 13 mutations of various fitness effects at 6 positions, the presented landscape is not only uniquely large but also distinguishes itself from previously published work regarding several other experimental features—namely, by its systematic and controlled experimental setup using engineered mutations of various selective effects and by considering multiple alleles simultaneously. We begin by describing the landscape and identifying the global peak, which is reached through a highly positively epistatic combination of four mutations. Based on a variety of implemented statistical measures and models, we describe the accessibility of the peak, the pattern of epistasis, and the topography of the landscape. To accommodate our data, we extend several previously used models and statistics to the multiallelic case. Using subsets of the landscape, we discuss the predictive potential of such modeling and the problem of selecting nonrandom mutations when attempting to quantify local landscapes to extrapolate global features.

Results and Discussion

We used the EMPIRIC approach (24, 25) to assess the growth rate of 640 mutants in yeast Hsp90 (*Materials and Methods*). Based on previous screenings of fitness effects in different environments (26) and on different genetic backgrounds (20), and on expectations of their biophysical role, 13 amino acid-changing point mutations at 6 sites were chosen for the fitness landscape presented here (Fig. 1). The fitness landscape was created by assessing the growth rate associated with each individual mutation on the parental background and all possible mutational combinations. A previously described Monte Carlo Markov chain (MCMC) approach was used to assess fitness and credibility intervals (ref. 27 and *Materials and Methods*).

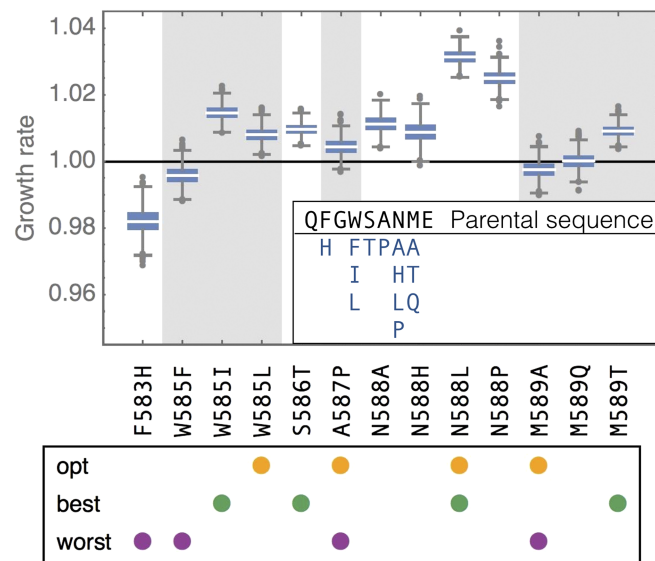


Fig. 1. Individual amino acid substitutions and their effect on the parental background in elevated salinity, obtained from 1,000 samples from the posterior distribution of growth rates. Boxes represent the interquartile range [i.e., the 50% confidence interval (CI)], whiskers extend to the highest/lowest data point within the box ± 1.5 times the interquartile range, and circles represent outliers; gray and white background shading alternates by amino acid position. The box below indicates with colored dots which mutations are involved in the focal landscapes discussed throughout the main text: the four mutations leading from the parental type to the global optimum ("opt"), the four individually most beneficial mutations on the parental background ("best"), and the four mutations with the individually lowest growth rates on the parental background ("worst"). (*Inset*) Parental sequence at positions 582–590 and assessed amino acids by position.

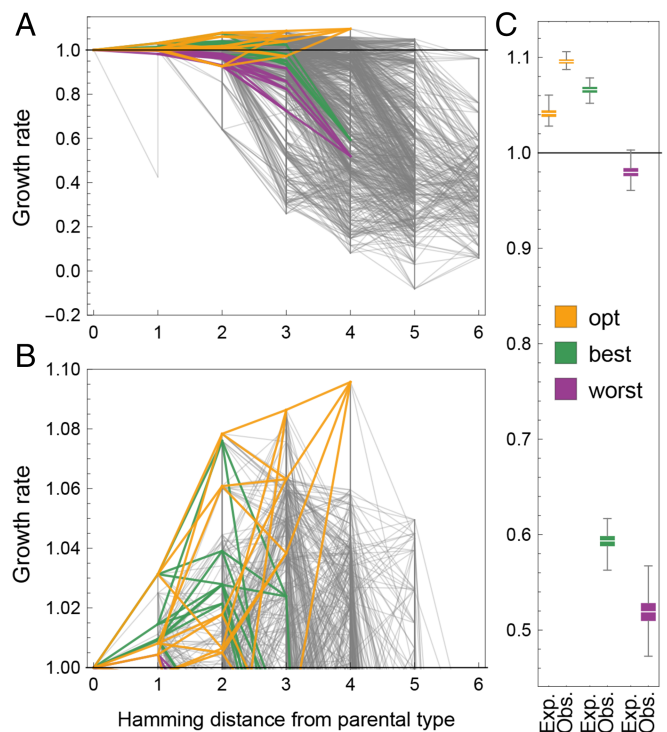


Fig. 2. (A) Empirical fitness landscape by mutational distance to parental type. Each line represents a single substitution. Vertical lines appear when multiple alleles have been screened at the same position. There is a global pattern of negative epistasis. The three focal landscapes are highlighted. (B) Close-up on the beneficial portion of the landscape. (C) Expected ("Exp.") versus observed ("Obs.") fitness for the focal landscapes, obtained from 1,000 posterior samples. We observe strong positive epistasis in the landscape that contains the global optimum, whereas the other two are dominated by negative epistasis. In A–C, the y axis depicts growth rate as a proxy for fitness.

The Fitness Landscape and Its Global Peak. Fig. 2 presents the resulting fitness landscape, with each mutant represented based on its Hamming distance from the parental genotype and its median estimated growth rate. Lines connect single-step substitutions, with vertical lines occurring when there are multiple mutations at the same position (Fig. 1). With increasing Hamming distance from the parental type, many mutational combinations become strongly deleterious. Thus, we observe strong negative epistasis between the substitutions that, as single steps on the parental background, have small effects. This pattern is consistent with Fisher's Geometric Model (28) when combinations of individually beneficial or small-effect mutations overshoot the optimum and with classic arguments predicting negative epistasis based on mutational load (29). It is also intuitively comprehensible on the protein level, where the accumulation of too many mutations is likely to destabilize the protein and render it dysfunctional (30).

The global peak of the fitness landscape is located four mutational steps away from the parental type (Fig. 2B), with 98% of posterior samples identifying the peak. The fitness advantage of the global peak reaches nearly 10% over the parental type and is consistent between replicates (*Materials and Methods* and *SI Appendix*, Fig. S3.1). Perhaps surprising given the degree of conservation of the studied genomic region (figure S5 in ref. 24), it is important to note that these fitness effects are measured under highly artificial experimental conditions including high salinity, which are unlikely to represent a natural environment of yeast. The effects of the individual mutations comprising the peak in a previous experiment without

added NaCl were -0.04135 , -0.01876 , -0.03816 , and -0.02115 for mutations W585L, A587P, N588L, and M589A, respectively, emphasizing the potential cost of adaptation associated with the increased salinity environment (data from ref. 20; see also ref. 26).

Curiously, the global peak is not reached by combining the most beneficial single-step mutations on the parental background but via a highly synergistic combination of one beneficial and three “neutral” mutations (i.e., mutations that are individually indistinguishable from the parental type in terms of growth rate). In fact, each of the five beneficial local optima shows a similar signature of positive epistasis (*SI Appendix, Fig. S3.2*). Fig. 2C demonstrates that a combination of the four mutations involved in the global peak (opt) predicts only a 4% fitness advantage. Furthermore, even a combination of the four individually most-beneficial, single-step mutations on the parental background in the dataset (best) (considering at most one mutation per position) only predicts a benefit of 6%. Notably, the actual combination of these four mutations on the parental background is highly deleterious and thus exhibits strong negative epistasis. Although negative epistasis between beneficial mutations during adaptation has been reported more frequently, positive epistasis has also been observed occasionally (18, 31), particularly in the context of compensatory evolution. In fact, negative epistasis between beneficial mutations and positive epistasis between neutral mutations has been predicted by de Visser et al. (32). Furthermore, our results support the pattern recently found in the gene underlying the antibiotic resistance enzyme TEM-1 β -lactamase in *Escherichia coli*, showing that large-effect mutations interact more strongly than small-effect mutations such that the fitness landscape of large-effect mutations tends to be more rugged than the landscape of small-effect mutations (13) and that mutations that were selected for their combined beneficial effect on a wild-type background tend to interact synergistically, whereas mutations selected for their individual effects interact antagonistically (4, 13).

Adaptive Walks on the Fitness Landscape. Next, we studied the empirical fitness landscape within a framework recently proposed by Draghi and Plotkin (33). Given the empirical landscape, we simulated adaptive walks and studied the accessibility of the six observed local optima. In addition, we evaluated the length of adaptive walks starting from any mutant in the landscape, until an optimum is reached. In the strong selection weak mutation limit (34), we can express the resulting dynamics as an absorbing Markov chain, where local optima correspond to the absorbing states, and in which the transition probabilities correspond to the relative fitness increases attainable by the neighboring mutations (*Materials and Methods*). This approach allowed us to derive analytical solutions for the mean and variance of the number of steps to reach a fitness optimum (*SI Appendix, Supporting Information 1: Extended Materials and Methods*), and the probability to reach a particular optimum starting from any given mutant in the landscape (Fig. 3 and *SI Appendix, Figs. S3.3 and S3.4*).

Using this framework, we find that the global optimum can be reached with nonzero probability from almost 95% of starting points in the landscape, and is reached with high probability from a majority of starting points - indicating high accessibility of the global optimum (Fig. 3). The picture changes when restricting the analysis to adaptive walks initiating from the parental type (Fig. 3 and *SI Appendix, Figs. S3.3 and S3.4*). Here, although 73% of all edges and 78% of all vertices are included in an adaptive walk to the global optimum, it is reached with only 26% probability. A local optimum two substitutions away from the parental type (Fig. 3C) is reached with a much higher probability of 47%. Hence, adaptation on the studied landscape is likely to stall at a suboptimal fitness peak. This observation indicates that the local and global landscape pattern may be quite different, an

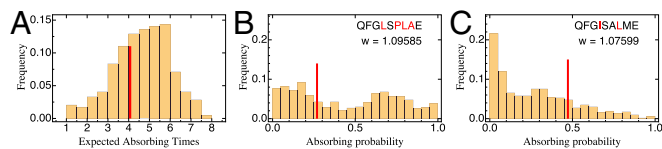


Fig. 3. (A) Distribution of average lengths of adaptive walks starting from any type in the full landscape (i.e., absorbing times of the Markov chain). The red line indicates mean path length for adaptive walks from the parental genotype. (B and C) Distribution of absorbing probabilities, that is, the probability to reach a specific optimum starting from a given genotype computed for all genotypes in the dataset. The red line corresponds to the respective probability when starting from the parental sequence. The global optimum (B) is, in general, reached with a very high probability, but there are starting points from which it is poorly accessible.

observation that is confirmed and discussed in more detail below (see *Predictive Potential of Landscape Statistics*). In line with the existence of multiple local fitness peaks, we find that pairs of alleles at different loci show pervasive sign (30%) and reciprocal sign (8%) epistasis (35), whereas the remaining 62% are attributed to magnitude epistasis (i.e., there is no purely additive interaction between alleles; for a discussion of the contribution of experimental error, see *SI Appendix, Fig. S3.5*).

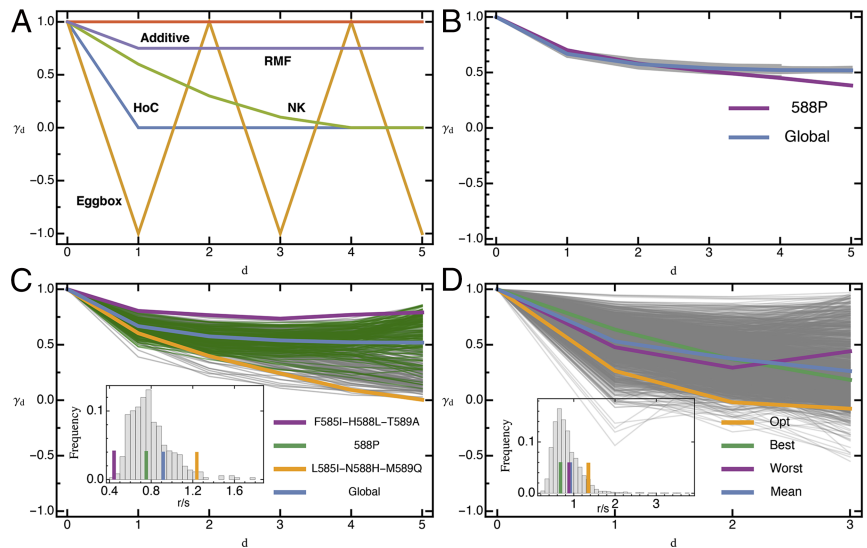
Epistasis Measures and the Topography of the Fitness Landscape.

Next, we considered the global topography of the fitness landscape. Various measures of epistasis and ruggedness have been proposed, most of them correlated and hence capturing similar features of the landscape (10). However, drawing conclusions has proven difficult because the studied landscapes were created according to different criteria. Furthermore, the majority of published complete landscapes are too small to be divided into subsets (but see refs. 10 and 12), preventing tests for the consistency and hence the predictive potential of landscape statistics. The landscape studied here provides us with this opportunity. Moreover, because multiple alleles at the same site are contained within the landscape, we may study whether changes in the shape of the landscape are site- or amino acid-specific.

We computed various landscape statistics (roughness-to-slope ratio, fraction of epistasis, and the recently proposed gamma statistics; *SI Appendix, Supporting Information 1: Extended Materials and Methods*) (10, 11, 14) and compared them with expectations from theoretical landscape models [NK (36–38), RMF (39, 40), HoC (41), egg-box landscapes (14); for brief definitions of these terms, see *SI Appendix, Supporting Information 2: Overview of Different Fitness Landscape Models Introduced in the Main Text*]. Whenever necessary, we provide an analytical extension of the used statistic to the case of multiallelic landscapes (*Materials and Methods*). To assess consistency and predictive potential, we computed the whole set of statistics for (i) all landscapes in which 1 amino acid was completely removed from the landscape (a cross-validation approach (42), subsequently referred to as the “drop-one” approach), (ii) all possible 360 diallelic sublandscapes, and (iii) for all 1,570 diallelic, 4-step landscapes containing the parental genotype, highlighting as special examples the three focal landscapes discussed.

We find that the general topography of the fitness landscape resembles that of a RMF landscape with intermediate ruggedness, which is characterized by a mixture of a random HoC component and an additive component (Fig. 4A and B). Whereas the whole set of landscape statistics supports this topography and our conclusions, the gamma statistics measuring landscape-averaged correlations in fitness effects, recently proposed in ref. 14, proved to be particularly illustrative. We will therefore focus on these in the main text; we refer to *SI Appendix* for additional results (e.g., measurement uncertainty and adaptive walks).

Fig. 4. (A) Expected pattern of landscape-wide epistasis measure γ_d (SI Appendix, Eq. S1.15) with mutational distance for theoretical fitness landscapes with 6 loci from ref. 14. (B) Observed decay of γ_d with mutational distance under the drop-one approach is quite homogenous, except when hot spot mutation 588P is removed; 95% CIs are contained within the lines for the global landscape. (C) Observed decay of γ_d for all diallelic six-locus sublandscapes. Depending on the underlying mutations, γ_d is vastly different, suggesting qualitative differences in the topography of the underlying fitness landscape and in the extent of additivity in the landscape. Three focal landscapes representative of different types and γ_d for the full landscape have been highlighted. (D) Observed decay of γ_d for all diallelic four-locus sublandscapes containing the parental type, indicative of locally different landscape topographies. Highlighted are the three focal landscapes. (Insets) Histograms of the roughness-to-slope ratio r/s for the respective subset, with horizontal lines indicating values for highlighted landscapes [r/s for global landscape (blue) is significantly different from HoC expectation, $p = 0.01$]. Similar to γ_d , there is a huge variation in r/s for subparts of the fitness landscape, especially when considering the four-locus subsets.



Predictive Potential of Landscape Statistics. When computed based on the whole landscape and on a drop-one approach, the landscape appears quite homogeneous, and the gamma statistics show relatively little epistasis (Figs. 4B and 5B). At first sight, this result contradicts our earlier statement of strong negative and positive epistasis but can be understood, given the different definitions of the epistasis measures used: above, we have measured epistasis based on the deviation from the multiplicative combination of the single-step fitness effects of mutations on the parental background. Because these effects were small, epistasis was strong in comparison. Conversely, the gamma measure is independent of a reference genotype and captures the fitness decay with a growing number of substitutions as a dominant and quite additive component of the landscape.

Only mutation 588P has a pronounced effect on the global landscape statistics and seems to act as an epistatic hotspot by making a majority of subsequent mutations (of individually small effect) on its background strongly deleterious (clearly visible in Figs. 4B and 5C). This finding can be explained by looking at the biophysical properties of this mutation. In wild-type Hsp90, amino acid 588N is oriented away from solvent and forms hydrogen bond interactions with neighboring amino acids (24). Proline lacks an amide proton, which inhibits hydrogen bond interactions. As a result, substituting 588N with a proline could disrupt hydrogen bond interactions with residues that may be involved in main chain hydrogen bonding and destabilize the protein. In addition, the pyrrolidine ring of proline is extremely rigid and can constrain the main chain, which may restrict the conformation of the residue preceding it in the protein sequence (43).

The variation between inferred landscape topographies increases dramatically for the 360 diallelic, 6-locus sublandscapes (Fig. 4C). Whereas all sublandscapes are largely compatible with an RMF landscape, the decay of landscape-wide epistasis with mutational distance (as measured by γ_d) shows a large variance, suggesting large differences in the degree of additivity. Interestingly, various sublandscapes, typically carrying mutation *588P, show a relaxation of epistatic constraint with increasing mutational distance (i.e., increasing γ_d) that is not captured by any of the proposed theoretical fitness landscape models, suggestive of systematic compensatory interactions (but see the egg-box model for an explicit example featuring nonmonotonicity in γ_d). The variation in the shape of the fitness (sub)landscapes is also

reflected in the corresponding roughness-to-slope ratio (inset of Fig. 4C and D), further emphasizing heterogeneity of the fitness landscape with local epistatic hotspots.

Finally, the 1,570 di-allelic four-locus landscapes containing the parental genotype, although highly correlated genetically, reflect a variety of possible landscape topographies (Fig. 4D), ranging from almost additive to egg-box shapes, accompanied by an extensive range of roughness-to-slope ratios. The three focal landscapes discussed above are not strongly different compared with the overall variation and yet show diverse patterns of epistasis between substitutions (Fig. 5).

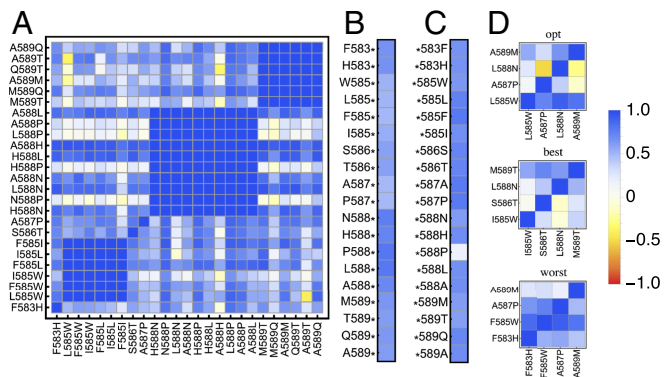


Fig. 5. (A) Epistasis measure $\gamma_{(A_i, B_i) \rightarrow (A_j, B_j)}$ (SI Appendix, Eq. S1.8; (A_i, B_i) on the x axis) between any two substitutions, averaged across the entire landscape. The majority of interactions are small to moderate (blue). Parts of the fitness landscape show highly localized and mutation-specific epistasis, ranging from strong magnitude epistasis (white) to sign and reciprocal sign epistasis (yellow). (B) The average epistatic effect $\gamma_{A_j \rightarrow}$ (SI Appendix, Eq. S1.10) of a mutation occurring on any background is always small. (C) The average epistatic effect $\gamma_{\rightarrow A_j}$ (SI Appendix, Eq. S1.13) of a background on any new mutation is usually small, except for mutations to *588P, which shows a strong magnitude effect. (D) Locus-specific gamma for the four mutations leading to the global optimum (Top), the four largest single-effect mutations (Middle), and the single-effect mutations with the lowest fitness (Bottom). The opt landscape exhibits strong sign epistasis between loci 587 and 588 and between loci 588 and 589. Also, the best landscape exhibits pervasive epistasis with sign epistasis between locus 588 and loci 585 and 586, respectively. We observe almost no epistasis in the worst landscape.

Thus, predicting fitness landscapes is difficult indeed. Extrapolation of the landscape, even across only a single mutation, may fail due to the existence of local epistatic hotspot mutations. Although the integration of biophysical properties into landscape models is an important step forward (e.g., ref. 44), we demonstrate that such models need to be mutation-specific. Considering a site-specific model [e.g., BLOSUM matrix (45)] is not sufficient. Newer models such as DeepAlign may provide the opportunity to allow integration of mutation-specific effects via aligning two protein structures based on spatial proximity of equivalent residues, evolutionary relationship, and hydrogen bonding similarity (46).

Conclusion

Originally introduced as a metaphor to describe adaptive evolution, fitness landscapes promise to become a powerful tool in biology to address complex questions regarding the predictability of evolution and the prevalence of epistasis within and between genomic regions. Due to the high-dimensional nature of fitness landscapes, however, the ability to extrapolate will be paramount to progress in this area, and the optimal quantitative and qualitative approaches to achieve this goal are yet to be determined.

Here, we have taken an important step toward addressing this question via the creation and analysis of a landscape comprising 640 engineered mutants of the Hsp90 protein in yeast. The unprecedented size of the fitness landscape, along with the multiallelic nature, allows us to test whether global features could be extrapolated from subsets of the data. Although the global pattern indicates a rather homogeneous landscape, smaller sublandscapes are a poor predictor of the overall global pattern because of “epistatic hotspots.”

In combination, our results highlight the inherent difficulty imposed by the duality of epistasis for predicting evolution. In the absence of epistasis (i.e., in a purely additive landscape), evolution is globally highly predictable because the population will eventually reach the single-fitness optimum, but the path taken is locally entirely unpredictable. Conversely, in the presence of (sign and reciprocal sign) epistasis evolution is globally unpredictable, because there are multiple optima and the probability to reach any one of them depends strongly on the starting genotype. At the same time, evolution may become locally predictable with the population following obligatory adaptive paths that are a direct result of the creation of fitness valleys owing to epistatic interactions.

The empirical fitness landscape studied here appears to be intermediate between these extremes. Although the global peak is within reach from almost any starting point, there is a local optimum that will be reached with appreciable probability, particular when starting from the parental genotype. From a practical standpoint, these results thus highlight the danger inherent to the common practice of constructing fitness landscapes from ascertained mutational combinations. However, this work also suggests that one promising way forward for increasing predictive power will be the utilization of multiple small landscapes used to gather information about the properties of individual mutations, combined with the integration of site-specific biophysical properties.

Materials and Methods

Here, we briefly outline the materials and methods used. A more detailed treatment of the theoretical work is presented in *SI Appendix*.

Data Generation. Codon substitution libraries consisting of 640 combinations (single up to sextuplet mutants) of 13 previously isolated individual mutants within the 582-to-590 region of yeast Hsp90 were generated from optimized cassette-ligation strategies, as previously described (24) and cloned into the p417GPD plasmid that constitutively expresses Hsp90.

Constitutively expressed libraries of Hsp90 mutation combinations were introduced into the *S. cerevisiae* shutoff strain DBY288 (can1-100 ade2-1 his3-11,15 leu2-3,12 trp1-1 ura3-1 hsp82: :leu2 hsc82: :leu2 ho: :pgals-hsp82-his3) using the lithium acetate method (47). Following transformation, the library was amplified for 12 h at 30 °C under nonselective conditions using galactose (Gal) medium with 100 μ g/mL ampicillin (1.7 g/L yeast nitrogen base without amino acids, 5 g/L ammonium sulfate, 0.1 g/L aspartic acid, 0.02 g/L arginine, 0.03 g/L valine, 0.1 g/L glutamic acid, 0.4 g/L serine, 0.2 g/L threonine, 0.03 g/L isoleucine, 0.05 g/L phenylalanine, 0.03 g/L tyrosine, 0.04 g/L adenine hemisulfate, 0.02 g/L methionine, 0.1 g/L leucine, 0.03 g/L lysine, and 0.01 g/L uracil with 1% raffinose and 1% Gal). After amplification, the library culture was transferred to selective medium similar to Gal medium but raffinose and Gal are replaced with 2% (wt/vol) dextrose. The culture was grown for 8 h at 30 °C to allow shutoff of the wild-type copy of Hsp90 and then shifted to selective medium containing 0.5 M NaCl for 12 generations. Samples were taken at specific time points and stored at -80 °C.

Yeast lysis, DNA isolation, and preparation for Illumina sequencing were performed as previously described (25). Sequencing was performed by Elim Biopharmaceuticals and produced ≈ 30 million reads of 99% confidence at each read position based on PHRED scoring (48, 49). Analysis of sequencing data were performed as previously described (26).

Estimation of Growth Rates. Individual growth rates were estimated according to the approach described in ref. 20 using a Bayesian MCMC approach proposed in ref. 27. Nucleotide sequences coding for the same amino acid sequence were interpreted as replicates with equal growth rates. The resulting MCMC output consisted of 10,000 posterior estimates for each amino acid mutation corresponding to an average effective samples size of 7,419 (minimum 725). Convergence was assessed using the Hellinger distance approach (50) combined with visual inspection of the resulting trace files.

Adaptive Walks. In the strong selection weak mutation limit (51), adaptation can be modeled as a Markov process only consisting of subsequent fitness-increasing, one-step substitutions that continue until an optimum is reached (so-called adaptive walks). This process is characterized by an absorbing Markov chain with a total of n different states (i.e., mutants), consisting of k absorbing (i.e., optima) and $n - k$ transient states (i.e., nonoptima). Defining $w(g)$ as the fitness of genotype g , and $g_{[i]}$ as the genotype g carrying a mutant allele at locus i , the selection coefficient is denoted by $s_i(g) = w(g_{[i]}) - w(g)$, such that the transition probabilities $p_{g, g_{[i]}}$ for going from any mutant g to any mutant $g_{[i]}$ are given by the fixation probability (52, 53) normalized by the sum over all adaptive, one-mutant neighbors of the current genotype g . If g is a (local) optimum, $p_{g, g} = 1$. Putting the transition matrix \mathbf{P} (54) in its canonical form and computing the fundamental matrix then allows to determine the expectation and the variance in the number of steps before reaching any optimum and to calculate the probability to reach optimum g when starting from genotype g' (55). Robustness of the results and the influence of specific mutations were assessed by deleting the corresponding columns and rows in \mathbf{P} (i.e., by essentially treating the corresponding mutation as unobserved) and recalculating and comparing all statistics to those obtained from the full dataset.

Correlation of Fitness Effects of Mutations. Strength and type of epistasis were assessed by calculating the correlation of fitness effects of mutations γ (14), which quantifies how the selective effect of a focal mutation is altered when put onto a different genetic background, averaged over all genotypes of the fitness landscape. Extending recent theory (14), we calculated the matrix of epistatic effects between different pairs of alleles (A_i, B_j) and (A_j, B_i) termed $\gamma_{(A_i, B_i) \rightarrow (A_j, B_j)}$ (*SI Appendix, Eq. S1.8*), the vector of epistatic effects between a specific pair of alleles (A_i, B_i) on all other pairs of alleles $\gamma_{(A_i, B_i) \rightarrow}$ (*SI Appendix, Eq. S1.9*), the vector of epistatic effects between all pairs of alleles on a specific allele pair (A_j, B_j) termed $\gamma_{\rightarrow (A_j, B_j)}$ (*SI Appendix, Eq. S1.12*), and the decay of correlation of fitness effects γ_d (*SI Appendix, Eq. S1.15*) with Hamming distance d averaged over all genotypes g of the fitness landscape.

Fraction of Epistasis. Following refs. 35 and 56, we quantified whether specific pairs of alleles between two loci interact epistatically and, if so, whether these display magnitude epistasis [i.e., fitness effects are nonadditive (57), but fitness increases with the number of mutations], sign epistasis (i.e., one of the two mutations considered has an opposite effect in both backgrounds) or reciprocal sign epistasis (i.e., if both mutations show sign epistasis). In particular, we calculated the type of epistatic interaction between mutations $g_{[i]}$ and $g_{[j]}$ (with $i \neq j$) with respect to a given reference geno-

type g over the entire fitness landscape. There was no epistatic interaction if $|s_j(g_{[ij]} - s_j(g))| < \varepsilon = 10^{-6}$, magnitude epistasis if $s_j(g)s_j(g_{[ij]}) \geq 0$ and $s_j(g)s_j(g_{[ij]}) \geq 0$, reciprocal sign epistasis if $s_j(g)s_j(g_{[ij]}) < 0$ and $s_j(g)s_j(g_{[ij]}) < 0$, and sign epistasis in all other cases (58).

Roughness-to-Slope Ratio. Following ref. 11, we calculated the roughness-to-slope ratio ρ by fitting the fitness landscape to a multidimensional linear model using the least-squares method. The slope of the linear model corresponds to the average additive fitness effect (10, 23), whereas the roughness is given by the variance of the residuals. Generally, the better the linear model fit, the smaller the variance in residuals such that the roughness-to-slope ratio approaches 0 in a perfectly additive model. Conversely, a very rugged fitness landscape would have a large residual

variance and, thus, a very large roughness-to-slope ratio (as in the HoC model). In addition, we calculated a test statistic Z_ρ by randomly shuffling fitness values in the sequence space to evaluate the statistical significance of the obtained roughness-to-slope ratio from the dataset ρ_{data} given by $Z_\rho = \frac{\rho_{\text{data}} - E[\rho_{\text{RSL}}]}{\sqrt{\text{Var}[\rho_{\text{RSL}}]}}$.

ACKNOWLEDGMENTS. We thank Dan Bolon, Brian Charlesworth, Pamela Cote, Inês Fragata, and Hermina Ghenu for helpful comments and discussion. This project was funded by grants from the Swiss National Science Foundation and a European Research Council Starting Grant (to J.D.J.). Computations were performed at the Vital-IT Center (www.vital-it.ch) for high-performance computing of the Swiss Institute of Bioinformatics.

- Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, ed Jones DF (Brooklyn Botanic Garden, Menasha, WI), pp 356–366.
- Coyne JA, Barton NH, Turelli M (1997) Perspective: A critique of Sewall Wright's shifting balance theory of evolution. *Evolution* 51(3):643–671.
- Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. (Princeton Univ Press, Princeton).
- de Visser JAGM, Krug J (2014) Empirical fitness landscapes and the predictability of evolution. *Nat Rev Genet* 15(7):480–490.
- Costanzo M, et al. (2010) The genetic landscape of a cell. *Science* 327(5964):425–431.
- Jiménez JI, Xulvi-Brunet R, Campbell GW, Turk-Macleod R, Chen IA (2013) Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc Natl Acad Sci USA* 110(37):14984–14989.
- Huang S (2013) Genetic and non-genetic instability in tumor progression: Link between the fitness landscape and the epigenetic landscape of cancer cells. *Cancer Metastasis Rev* 32(3–4):423–448.
- Mann JK, et al. (2014) The fitness landscape of HIV-1 gag: Advanced modeling approaches and validation of model predictions by in vitro testing. *PLoS Comput Biol* 10(8):e1003776.
- Liu G, Rancati G (2016) Adaptive evolution: Don't fix what's broken. *Curr Biol* 26(4):R169–R171.
- Szendro IG, Schenk MF, Franke J, Krug J (2013) Quantitative analyses of empirical fitness landscapes. *J Stat Mech* 2013:P01005.
- Aita T, Iwakura M, Husimi Y (2001) A cross-section of the fitness landscape of dihydrofolate reductase. *Protein Eng* 14(9):633–638.
- Franke J, Klözer A, de Visser JAGM, Krug J (2011) Evolutionary accessibility of mutational pathways. *PLoS Comput Biol* 7(8):e1002134.
- Schenk MF, Szendro IG, Salverda MLM, Krug J, de Visser JAGM (2013) Patterns of Epistasis between beneficial mutations in an antibiotic resistance gene. *Mol Biol Evol* 30(8):1779–1787.
- Ferretti L, et al. (2016) Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. *J Theor Biol* 396:132–143.
- Kryazhinskiy S, Rice DP, Jerison ER, Desai MM (2014) Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344(6191):1519–1522.
- Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332(6034):1193–1196.
- Schoustra SE, Bataillon T, Gifford DR, Kassen R (2009) The properties of adaptive walks in evolving populations of fungus. *PLoS Biol* 7(11):e1000250.
- Weinreich D, Delaney N, DePristo M (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770):111–114.
- Dickinson WJ (2008) Synergistic fitness interactions and a high frequency of beneficial changes among mutations accumulated under relaxed selection in *Saccharomyces cerevisiae*. *Genetics* 178(3):1571–1578.
- Bank C, Hietpas RT, Jensen JD, Bolon DNA (2015) A systematic survey of an intragenic epistatic landscape. *Mol Biol Evol* 32(1):229–238.
- MacLean RC, Hall AR, Perron GG, Buckling A (2010) The population genetics of antibiotic resistance: Integrating molecular mechanisms and treatment contexts. *Nat Rev Genet* 11(6):405–414.
- Jasmin JN, Lenormand T (2016) Accelerating mutational load is not due to Synergistic epistasis or mutator alleles in mutation accumulation lines of yeast. *Genetics* 202(2):751–763.
- Blanquart F, Achaz G, Bataillon T, Tenaillon O (2014) Properties of selected mutations and genotypic landscapes under Fisher's geometric model. *Evolution* 68(12):3537–3554.
- Hietpas RT, Jensen JD, Bolon DNA (2011) Experimental illumination of a fitness landscape. *Proc Natl Acad Sci USA* 108(19):7896–7901.
- Hietpas R, Roscoe B, Jiang L, Bolon DNA (2012) Fitness analyses of all possible point mutations for regions of genes in yeast. *Nat Protoc* 7(7):1382–1396.
- Hietpas RT, Bank C, Jensen JD, Bolon DNA (2013) Shifting fitness landscapes in response to altered environments. *Evolution* 67(12):3512–3522.
- Bank C, Hietpas RT, Wong A, Bolon DN, Jensen JD (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.
- Weinreich DM, Knies JL (2013) Fisher's geometric model of adaptation meets the functional synthesis: Data on pairwise epistasis for fitness yields insights into the shape and size of phenotype space. *Evolution* 67(10):2957–2972.
- Kimura M, Maruyama T (1966) Mutational load with epistatic gene interactions in fitness. *Genetics* 54(6):1337–1351.
- Soskine M, Tawfik DS (2010) Mutational effects and the evolution of new protein functions. *Nat Rev Genet* 11(8):572–582.
- Trindade S, et al. (2009) Positive epistasis drives the acquisition of multidrug resistance. *PLoS Genet* 5(7):e1000578.
- de Visser JA, Hoekstra RF, van den Ende H (1997) An experimental test for synergistic epistasis and its application in *Chlamydomonas*. *Genetics* 145(3):815–819.
- Draghi JA, Plotkin JB (2013) Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution* 67(11):3120–3131.
- Gillespie JH (1984) Molecular evolution over the mutational landscape. *Evolution* 38(5):1116–1129.
- Weinreich DM, Watson RA, Chao L (2005) Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59(6):1165–1174.
- Kauffman S, Levin S (1987) Towards a general theory of adaptive walks on rugged landscapes. *J Theor Biol* 128(1):11–45.
- Kauffman SA (1993) *The Origins of Order: Self Organization and Selection in Evolution*. (Oxford Univ Press, New York).
- Schmiegel B, Krug J (2014) Evolutionary accessibility of modular fitness landscapes. *J Stat Phys* 154(1):334–355.
- Neidhart J, Szendro IG, Krug J (2014) Adaptation in tunably rugged fitness landscapes: The Rough Mount Fuji model. *Genetics* 198(2):699–721.
- Aita T, et al. (2000) Analysis of a local fitness landscape with a model of the rough Mt. Fuji-type landscape: Application to prolyl endopeptidase and thermolysin. *Biopolymers* 54(1):64–79.
- Kingman JFC (1978) A simple model for the balance between selection and mutation. *J Appl Probab* 15(1):1–12.
- Geisser S (1993) *Predictive Inference*. (CRC Press, New York), Vol 55.
- Bajaj K, et al. (2007) Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput Biol* 3(12):e241.
- Wylie CS, Shakhnovich EI (2011) A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci USA* 108(24):9916–9921.
- Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89(22):10915–10919.
- Wang S, Ma J, Peng J, Xu J (2013) Protein structure alignment beyond spatial proximity. *Sci Rep* 3:1448.
- Gietz DR, Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350:87–96.
- Ewing B, Green P (1998) Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* 8:186–194.
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* 8:175–185.
- Boone EL, Merrick JRW, Krachey MJ (2012) A Hellinger distance approach to MCMC diagnostics. *J Stat Comput Simulat* 84(4):833–849.
- Gillespie JH (1983) A simple stochastic gene substitution model. *Theor Popul Biol* 23(2):202–215.
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. (Clarendon Press, Oxford, UK).
- Wright S (1931) Evolution in Mendelian populations. *Genetics* 16(2):97.
- McCandlish DM (2011) Visualizing fitness landscapes. *Evolution* 65(6):1544–1558.
- Kemery JG, Snell JL (1960) Finite Markov Chains (van Nostrand, Princeton).
- Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445(7126):383–386.
- Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinburgh* 52:399–433.
- Brouillet S, Annoni H, Ferretti L, Achaz G (2015) MAGELLAN: A tool to explore small fitness landscapes. bioRxiv:031583.