

Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein

Shimon Bershtein¹, Michal Segal¹, Roy Bekerman¹, Nobuhiko Tokuriki¹ & Dan S. Tawfik¹

The distribution of fitness effects of protein mutations is still unknown^{1,2}. Of particular interest is whether accumulating deleterious mutations interact, and how the resulting epistatic effects shape the protein's fitness landscape. Here we apply a model system in which bacterial fitness correlates with the enzymatic activity of *TEM-1* β -lactamase (antibiotic degradation). Subjecting *TEM-1* to random mutational drift and purifying selection (to purge deleterious mutations) produced changes in its fitness landscape indicative of negative epistasis; that is, the combined deleterious effects of mutations were, on average, larger than expected from the multiplication of their individual effects. As observed in computational systems^{3–5}, negative epistasis was tightly associated with higher tolerance to mutations (robustness). Thus, under a low selection pressure, a large fraction of mutations was initially tolerated (high robustness), but as mutations accumulated, their fitness toll increased, resulting in the observed negative epistasis. These findings, supported by FoldX stability computations of the mutational effects⁶, prompt a new model in which the mutational robustness (or neutrality) observed in proteins, and other biological systems, is due primarily to a stability margin, or threshold, that buffers the deleterious physico-chemical effects of mutations on fitness. Threshold robustness is inherently epistatic—once the stability threshold is exhausted, the deleterious effects of mutations become fully pronounced, thereby making proteins far less robust than generally assumed.

Interactions between mutations have been studied by a number of fields, although nomenclature differs; geneticists term interactions within, or between genes, as 'epistasis'⁷, whereas protein biophysicists use 'double mutant cycles'⁸ to describe intragenic interactions. Because the ultimate impact of mutations depends both on the effect of each mutation on fitness and on the interactions between accumulating mutations, insights from these different disciplines can nevertheless be combined in one model².

Taking the converse view, if deleterious mutations do not interact, then under no selection, fitness (or the stability, or activity, of a protein) should decline exponentially:

$$W_{(n)} = \exp(-\alpha n) \quad (1)$$

where n is the number of mutations and α is the exponential decline parameter^{3,9}. If, however, deleterious mutations interact so that their combined impact on fitness is greater than expected from multiplying their individual effects (or more than additive in terms of $\log W$, or $\Delta\Delta G$ for protein stability and function), fitness decline would accelerate with the accumulation of mutations, giving rise to 'negative epistasis':

$$W_{(n)} = \exp(-\alpha n - \beta n^2) \quad (2)$$

where α reflects the fraction of multiplicative deleterious mutations and β is the epistasis parameter⁹.

Because neither of these crucial factors (α , β) is quantitatively understood—particularly at the level of a single protein^{1,2}—we set up an experimental system that measured the fitness landscape of a protein subjected to a lengthy random drift (up to 20 mutations per gene). We derived the exponential decline and epistasis parameters and examined their impact on the rate and dynamics of mutation accumulation.

As our model protein we chose *TEM-1* β -lactamase, an enzyme that degrades the antibiotic ampicillin and thereby confers ampicillin resistance on Gram-negative bacteria such as *Escherichia coli*. The fitness of a *TEM-1* variant is directly correlated with the maximal concentration of ampicillin that *E. coli* carrying this gene variant can tolerate. At the level of a population, fitness (W) refers to the fraction of variants that confer resistance at a given concentration of ampicillin. By determining the fraction of viable variants over the entire range of ampicillin concentrations, a fitness landscape was obtained (Supplementary Fig. 1).

The *TEM-1* gene was cloned into a plasmid (as it occurs in nature) under its endogenous promoter. Recloning after each round of mutagenesis confined the mutational drift to the open reading frame of *TEM-1*. Our *in vitro* random mutagenesis protocol was optimized for high reproducibility and was calibrated to obtain, on average, two mutations per gene per round of mutagenesis. We maintained three populations of randomly drifting *TEM-1* genes: one population under no selection (*Lib0*), and the rest under purifying selection at 'high' and 'low' stringencies. Each population, or plasmid library, was separately mutated, ligated into an empty vector and transformed into *E. coli* host cells; it then underwent purifying selection: 'high' selection pressure (250 $\mu\text{g ml}^{-1}$ ampicillin; *Lib250*; Supplementary Fig. 2), and 'low' selection pressure (12.5 $\mu\text{g ml}^{-1}$ ampicillin; *Lib12.5*). After growth on selection plates, plasmid DNA was extracted from the surviving *E. coli* colonies, and the *TEM-1* genes were subjected to the next round of mutagenesis. Altogether, ten successive rounds of mutagenesis and purifying selection were performed. Loss of diversity was less than 50% per round, and a diversity of at least 10^6 variants per library was maintained throughout.

As expected, a rapid fitness decline was observed in *Lib0* (no selection). The fitness of the selected populations (*Lib12.5* and *Lib250*) remained unchanged under the threshold of selection, and decreased above that threshold (Supplementary Fig. 3). Fitness values (W) of the randomly drifting population (*Lib0*) were plotted as a function of the mutational load n . These plots showed consistent and significant deviations from simple exponential decays. This was particularly true at low fitness levels; indeed, at less than 1,000 $\mu\text{g ml}^{-1}$ ampicillin the data could only be fitted reliably to equation (2) (Fig. 1). The 'net' effect of deleterious mutations accumulating randomly in the

¹Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel.

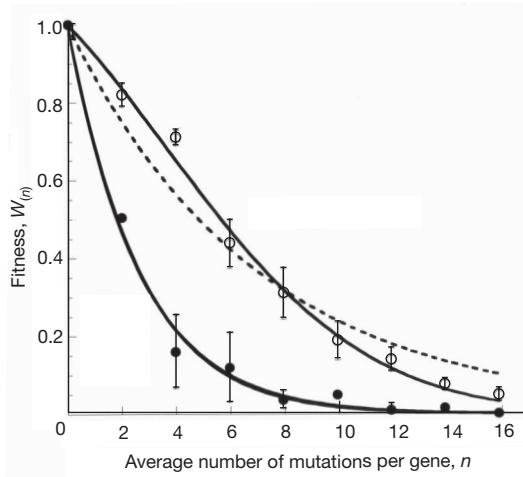


Figure 1 | Negative epistasis underlines the random drift of *TEM-1*. The fitness decline of unselected library *Lib0* at the highest fitness level (2,500 $\mu\text{g ml}^{-1}$ ampicillin; filled circles) is exponential and fits equation (1) with $\alpha = 0.371$ (lower black line). At the lowest ampicillin concentration (12.5 $\mu\text{g ml}^{-1}$; open circles) the data fit equation (2) well with $\alpha = 0.072$ and $\beta = 0.009$ (upper black line), but poorly to equation (1) ($\alpha = 0.145$; broken line). Error bars represent s.d. for two to five independent measurements of fitness.

TEM-1 gene was therefore found to be synergistic ($\beta > 0$; negative epistasis).

The α and β parameters deduced from fitness decay plots for all ampicillin concentrations (Supplementary Fig. 4) are plotted in Fig. 2a. The exponential decay parameter of *TEM-1* (α) increased almost linearly with fitness level. Extrapolating α to zero ampicillin indicated that the fraction of *TEM-1* mutations that are unconditionally lethal is about 7% (Fig. 2a), corresponding to mutations that severely undermine the stability of *TEM-1* or abolish its catalytic activity. The degree of negative epistasis was tightly correlated with fitness levels; a higher decline in fitness (larger α values) gave rise to weaker epistasis (β/α) (Fig. 2b). It therefore seems that when high fitness levels are maintained, the fraction of lethal mutations is also high (large α , low robustness). However, the remaining fraction of mutations is largely neutral, and their epistatic potential is minimal ($\beta/\alpha \approx 0$). Conversely, under low fitness levels a much larger fraction

of mutations is tolerated (small α , high robustness), but these exhibit large negative epistatic effects ($\beta/\alpha \gg 0$).

In terms of protein stability and function ($\Delta\Delta G$), the negative epistasis observed here implies that *TEM-1* mutations, on average, interact in a more-than-additive manner. However, this result does contradict our current knowledge of proteins. 'Double mutant cycles'⁶ is a commonly used tool to dissect proteins, where the effects of different mutations on various physico-chemical properties (thermodynamic stability, binding affinity or enzymatic rates) are measured on their own and together⁸. They therefore comprise a measure of epistasis. Numerous cycles performed on many different proteins indicate that the physico-chemical effects of deleterious mutations (in terms of $\Delta\Delta G$, or $\log W$) are mostly additive, or less than additive for interacting residues. Effects that are more than additive are rare exceptions⁸. We therefore surmise that the negative epistasis observed here is the outcome of a non-additive decline in fitness, in response to additive physico-chemical changes (see also ref. 2).

To examine this hypothesis we assumed that most deleterious mutations had undermined stability, thereby reducing the levels of soluble, active protein^{2,10}. We then applied the FoldX⁶ algorithm to predict the stability changes induced by mutations in the drifting *TEM-1* populations. FoldX analysis indicated that the destabilizing mutations frequently observed in the unselected library *Lib0* were purged by the purifying selections (Fig. 3a). Initially (as indicated by sampling of the fifth round), mutations with considerable destabilizing effects ($\Delta\Delta G \leq 6.5 \text{ kcal mol}^{-1}$) could accumulate in *Lib12.5*, whereas in *Lib250* only mutations exhibiting $\Delta\Delta G \leq 3.5 \text{ kcal mol}^{-1}$ were tolerated (a higher $\Delta\Delta G$ value corresponds to more destabilizing mutations). However, by the tenth round, only weakly destabilizing mutations ($\Delta\Delta G \leq 3.5 \text{ kcal mol}^{-1}$) accumulated in both populations (Fig. 3b).

The FoldX analysis therefore indicated the existence of a 'neutral' region in which changes in stability had only a mild effect on fitness, and, notably, that the width of this region changed with fitness level. As more mutations accumulated, this threshold was exhausted, and the tolerance to mutations under both regimes became essentially the same (Fig. 3b). This analysis is also in line with the finding that the percentage of tolerated mutations within *Lib12.5* was high at the beginning, and decreased during subsequent rounds (Table 1 and Supplementary Fig. 5). By the fifth round of *Lib12.5*, an average of 3.6 non-synonymous mutations per gene had accumulated, whereas only 1.9 mutations were added over the subsequent five rounds. In contrast, *Lib250* proceeded at a steady rate (2.2 non-synonymous

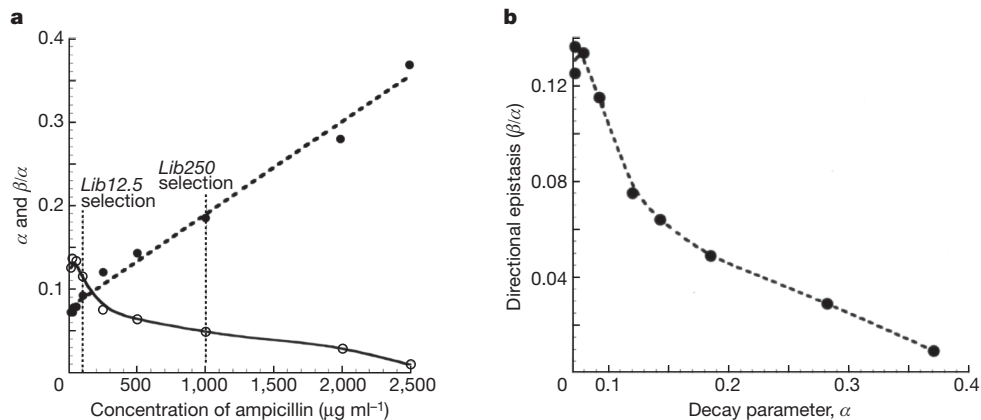


Figure 2 | The correlation between mutational robustness and negative epistasis. **a**, The decay parameter α (filled circles) and directional epistasis β/α (open circles) were extracted from the fitness decline fits of *Lib0* for each ampicillin concentration (Fig. 1, and Supplementary Fig. 4). These parameters are presented as a function of ampicillin concentration, corresponding to the fitness levels in our system. The vertical dotted lines

show the fitness levels for the purifying selections. **b**, The correlation between mutational robustness (higher α values indicate lower tolerance to mutations, and hence lower robustness) and epistasis. The same correlation was observed when an alternative measure of epistasis³ was applied (Supplementary Fig. 7).

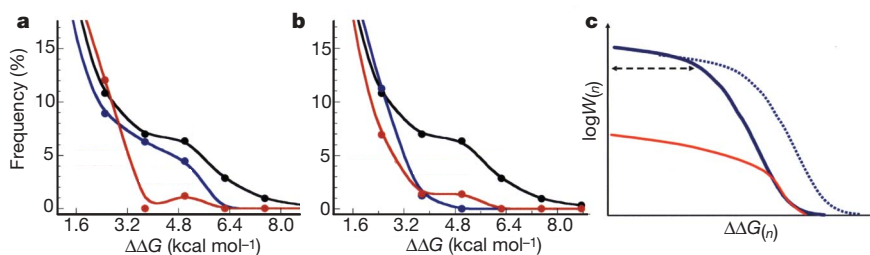


Figure 3 | The physico-chemical and fitness changes accompanying random drifts. **a, b,** The stability changes induced by 980 mutations identified in the three drifting populations were individually computed with FoldX. The calculated $\Delta\Delta G$ values were arranged in 1 kcal mol⁻¹ bins. Plotted are the frequencies of all destabilizing mutations found in the fifth round (**a**) and tenth round (**b**) for the unselected library (*Lib0*, black), and the two selected libraries (*Lib12.5* (blue) and *Lib250* (red)). **c,** Our model surmises that $\Delta\Delta G$ (physico-chemical changes relative to wild type, and in particular stability) declines linearly with the number of mutations

($\Delta\Delta G_{(n)} \propto n$). However, the decline in fitness $W_{(n)}$ that accompanies these $\Delta\Delta G$ changes is nonlinear and comprises both a threshold (dashed two-headed arrow) that buffers the deleterious effects of mutations, and a gradient in which fitness declines concomitantly with $\Delta\Delta G$ changes. At high fitness levels (solid blue line) the threshold is relatively narrow. At low fitness levels (red line) the threshold widens, giving rise to higher epistasis. A wider threshold is also predicted for protein variants carrying stabilizing mutations or global suppressors (dotted line; Supplementary Fig. 6).

mutations at the fifth round, and 4.6 at the tenth round). Indeed, the average numbers of mutations found in *Lib12.5* and *Lib250* by the tenth round of selection were nearly identical, despite a 20-fold difference in the stringency of selection.

It therefore seems that mutational robustness (or neutrality) should be described with two terms: ‘threshold’ and ‘gradient’ (Fig. 3c).

A ‘threshold’ induces a delay in fitness decline in the face of mutations. The threshold is observed not because the protein is unaffected by mutation but because the decline in its thermodynamic stability is buffered and has no immediate effect on fitness¹⁰. As mutations accumulate, however, the threshold is largely exhausted, and a gradient phase appears in which the fitness declines in parallel with $\Delta\Delta G$ changes. Because the robustness observed under lower fitness levels is largely due to a wider threshold, it is also inevitably epistatic (the fitness toll of later mutations is higher than that of the early ones). In fact, the manner in which α (robustness) and β (negative epistasis) correlate in our experiments (Fig. 2b) is in striking agreement with both theory¹¹ and simulations^{3–5}. The origins of our observed negative epistasis, and its correlation with robustness, therefore suggest that ‘threshold robustness’ is a general phenomenon that goes well beyond *TEM-1*.

Two conclusions can be derived from our results and model.

First, proteins may not be as robust as is generally assumed. So far, experiments have only tested the response to several random mutations and have reported an exponential decline in activity with no epistasis^{10,12,13}. The tolerance to mutations observed in these experiments is therefore likely to contain a major component of threshold robustness, especially as they were performed at low fitness^{10,14–16}. We, too, observed high levels of mutation tolerance under low fitness levels. Indeed, over the first five rounds of mutagenesis (70% of total mutations and 55% of non-synonymous mutations; Table 1), only 7% of mutations were unconditionally deleterious. These figures are

in agreement with the common view that the vast majority of protein mutations are tolerated^{10,14–17}. Yet once the stability threshold is exhausted, tolerance to mutations under the gradient regime is markedly lower and fits the view that most mutations affect fitness¹.

Second, our model provides general insights into the origins of robustness and epistasis, and why the two are interlinked^{3–5,11}. Threshold robustness relates to the margin of excess stability and function that could be compromised with little or no immediate effect on fitness, and is therefore inherently epistatic. Indeed, a quantitative correlation between the threshold (in $\Delta\Delta G$ terms) and directional epistasis (β/α) has been observed in our system and will be described in future work. Threshold robustness would be the type of robustness derived from redundancy at the genome level (for example duplicated genes)¹⁸, from stabilizing mutations¹⁹ or from global suppressors in single genes¹⁰ (Fig. 3c). It has been shown, for example, that *TEM-1* carrying a global suppressor mutation that increased stability by 2.7 kcal mol⁻¹ shows higher robustness relative to the wild type¹⁰ ($\alpha = 0.042$ versus 0.152 for wild type, by equation (1)). However, when we applied our model (equation (2)), we also recorded a much higher epistasis ($\beta/\alpha = 0.598$ versus 0.144; Supplementary Fig. 6).

Thus, theory and simulations^{3–5} have predicted a tight correlation between robustness and epistasis. Our work provides an experimental verification of this correlation and proposes a mechanism that accounts for it. Our model implies that any biological system that exhibits threshold robustness, or redundancy robustness¹⁸, is inevitably epistatic. In such systems, mechanisms that purge potentially deleterious mutations, such as recombination (through sexual reproduction and other mechanisms^{20,21}) are of crucial importance, as they help to maintain this threshold. In this way, recombination, threshold robustness and negative epistasis may be interlinked—each being an inevitable by-product of the other⁴.

Table 1 | Sequence data of the *TEM-1* libraries

Library	Round of mutagenesis and selection	No. of variants (base pairs) sequenced	Total no. of mutations detected	Average no. of mutations per gene	Average no. of non-synonymous mutations per gene	Percentage of tolerated mutations	
						Total	Non-synonymous
<i>Lib0</i>	1	10 (8,610)	16	1.6 ± 1.4	1.11 ± 1.0		
	2	21 (18,081)	81	3.9 ± 1.6	2.67 ± 1.1		
	3	5 (4,305)	25	5.0 ± 0.7	3.46 ± 0.5		
	5	17 (14,637)	161	9.5 ± 3.7	6.55 ± 2.6		
	10	14 (12,054)	290	20.7 ± 2.8	14.3 ± 1.9		
<i>Lib12.5</i>	2	16 (13,776)	52	3.3 ± 1.6	1.6 ± 1.2	84%	61%
	5	39 (33,579)	255	6.6 ± 3.0	3.6 ± 1.7	70%	55%
<i>Lib250</i>	10	17 (14,637)	171	10.7 ± 3.5	5.5 ± 2.2	52%	39%
	2	14 (12,054)	24	1.7 ± 1.2	1.3 ± 0.9	44%	48%
	5	39 (33,579)	193	5.0 ± 2.2	2.2 ± 1.5	52%	33%
	10	18 (15,498)	170	9.4 ± 3.0	4.6 ± 2.2	46%	32%

Where errors are shown, results are means ± s.d.

METHODS

Detailed methods are provided in Supplementary Information.

Library construction and selection. The *TEM-1* β -lactamase open reading frame was recloned (*NcoI/NotI*) under the control of its endogenous promoter into a modified pUC19 plasmid containing a chloramphenicol resistance gene. Random mutagenesis was performed by 'wobble'-base polymerase chain reaction with *TEM-1* plasmids as template. The protocol was optimized to achieve an average of two mutations per gene per generation, and a characteristic pattern of nucleotide exchange (Supplementary Table 2). Plasmid libraries were obtained by electroporation into *E. coli* and more than 10^6 individual transformants were grown in Luria-Bertani medium in the presence of chloramphenicol. DNA extracted from these cultures was retransformed into XL-1 Blue *E. coli* cells and plated on agar plates supplemented with ampicillin (0, 12.5 or 250 $\mu\text{g ml}^{-1}$). Plasmid DNA from surviving colonies was extracted and served as a template for the subsequent round of mutagenesis.

Fitness measurements. Several hundred colonies from each library were randomly picked and grown in 96-well plates in the presence of chloramphenicol. The 96-well plates were replica-plated onto agar plates with the entire range of ampicillin concentrations (0–2,500 $\mu\text{g ml}^{-1}$). Fitness values (W) were determined from the fraction of clones that survived each ampicillin concentration (two to five independent repetitions of this protocol were performed and standard deviations are presented). The reproducibility of the mutagenesis protocol was verified by an independent repetition of one round of mutagenesis, indicating a variability within the deviation of the fitness measurements (Supplementary Table 1).

Computation of stability effects. FoldX is a structure-based algorithm that predicts the $\Delta\Delta G$ values of mutations with considerable fidelity ($R = 0.83$ for a data set of 1,030 mutations in 27 proteins)⁶. We applied the published procedure, including an adjustment of the calculated values²². A correlation between the deleterious effect of mutations and their predicted $\Delta\Delta G$ values was also obtained (Supplementary Table 3).

Received 1 September; accepted 27 October 2006.

Published online 19 November 2006.

1. Pal, C., Papp, B. & Lercher, M. J. An integrated view of protein evolution. *Nature Rev. Genet.* **7**, 337–348 (2006).
2. DePristo, M. A., Weinreich, D. M. & Hartl, D. L. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature Rev. Genet.* **6**, 678–687 (2005).
3. Wilke, C. O. & Adami, C. Interaction between directional epistasis and average mutational effects. *Proc. R. Soc. Lond. B* **268**, 1469–1474 (2001).
4. Azevedo, R. B., Lohaus, R., Srinivasan, S., Dang, K. K. & Burch, C. L. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature* **440**, 87–90 (2006).
5. Elena, S. F., Carrasco, P., Daros, J. A. & Sanjuan, R. Mechanisms of genetic robustness in RNA viruses. *EMBO Rep.* **7**, 168–173 (2006).
6. Guerois, R., Nielsen, J. E. & Serrano, L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387 (2002).

7. Wolf, J. B., Brodie, E. D. & Wade, M. J. *Epistasis and the evolutionary process* (Oxford Univ. Press, Oxford, 2000).
8. Horovitz, A. Double-mutant cycles: a powerful tool for analyzing protein structure and function. *Fold. Des.* **1**, R121–R126 (1996).
9. Charlesworth, B. Mutation–selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**, 199–221 (1990).
10. Bloom, J. D. et al. Thermodynamic prediction of protein neutrality. *Proc. Natl Acad. Sci. USA* **102**, 606–611 (2005).
11. Wagner, G. P., Laubichler, M. D. & Bagheri-Chaichian, H. Genetic measurement of theory of epistatic effects. *Genetica* **102–103**, 569–580 (1998).
12. Shafikhani, S., Siegel, R. A., Ferrari, E. & Schellenberger, V. Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization. *Biotechniques* **23**, 304–310 (1997).
13. Daugherty, P. S., Chen, G., Iverson, B. L. & Georgiou, G. Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. *Proc. Natl Acad. Sci. USA* **97**, 2029–2034 (2000).
14. Guo, H. H., Choe, J. & Loeb, L. A. Protein tolerance to random amino acid change. *Proc. Natl Acad. Sci. USA* **101**, 9205–9210 (2004).
15. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. & Sauer, R. T. Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science* **247**, 1306–1310 (1990).
16. Sauer, U. H., San, D. P. & Matthews, B. W. Tolerance of T4 lysozyme to proline substitutions within the long interdomain α -helix illustrates the adaptability of proteins to potentially destabilizing lesions. *J. Biol. Chem.* **267**, 2393–2399 (1992).
17. Huang, W., Petrosino, J., Hirsch, M., Shenkin, P. S. & Palzkill, T. Amino acid sequence determinants of β -lactamase structure and activity. *J. Mol. Biol.* **258**, 688–703 (1996).
18. Wagner, A. Distributed robustness versus redundancy as causes of mutational robustness. *BioEssays* **27**, 176–188 (2005).
19. Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. A. Protein stability promotes evolvability. *Proc. Natl Acad. Sci. USA* **103**, 5869–5874 (2006).
20. Kondrashov, A. S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988).
21. Lynch, M., Burger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84**, 339–344 (1993).
22. Kiel, C. & Serrano, L. The ubiquitin domain superfold: structure-based sequence alignments and characterization of binding epitopes. *J. Mol. Biol.* **355**, 821–844 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Kondrashov for insights and inspiration, L. Serrano, U. Alon and G. Schreiber for contributions, and our Department, Institute and the Feinberg Graduate School, for uncompromising support.

Author Contributions S.B. designed, performed and analysed the experiments. M.S. and R.B. provided technical assistance. N.T. performed the FoldX computations and helped devise the model. D.S.T. designed and supervised the experiments, and devised the model. S.B. and D.S.T. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.S.T. (tawfik@weizmann.ac.il).