Title: Epistasis between promoter activity and coding mutations shapes gene evolvability

Authors: Angel F. Cisneros[1,2,3,4†], Isabelle Gagnon-Arsenault[1,2,3,4,5†], Alexandre K. Dubé[1,2,3,4,5], Philippe C. Després[1,2,3,4], Pradum Kumar[1,6], Kiana Lafontaine[3,7], Joelle N. Pelletier[3,7,8], Christian R. Landry[1,2,3,4,5*]

Affiliations:

[1]Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval; G1V 0A6, Québec, Canada

[2]Institut de biologie intégrative et des systèmes, Université Laval; G1V 0A6, Québec, Canada

[3]PROTEO, Le regroupement québécois de recherche sur la fonction, l'ingénierie et les applications des protéines, Université Laval; G1V 0A6, Québec, Canada

[4]Centre de recherche sur les données massives, Université Laval; G1V 0A6, Québec, Canada

[5]Département de biologie, Faculté des sciences et de génie, Université Laval; G1V 0A6, Québec, Canada

[6]Department of Biosciences and Bioengineering, Indian Institute of Technology Roorkee; 247667, Roorkee, India

[7]Département de biochimie et de médecine moléculaire, Faculté de médecine, Université de Montréal; H3C 3J7, Montréal, Canada

[8]Département de chimie, Faculté des arts et des sciences, Université de Montréal; H3C 3J7, Montréal, Canada

* Corresponding author. Email: christian.landry@bio.ulaval.ca

† These authors contributed equally to this work.

Abstract: The evolution of protein-coding genes proceeds as mutations act on two main dimensions: regulation of transcription level and the coding sequence. The extent and impact of the connection between these two dimensions are largely unknown because they have generally been studied independently. By measuring the fitness effects of all possible mutations on a protein complex at various levels of promoter activity, we show that promoter activity at the optimal level for the wild-type protein masks the effects of both deleterious and beneficial coding mutations. Mutations that are deleterious at low activity but masked at optimal activity are slightly destabilizing for individual subunits and binding interfaces. Coding mutations that increase protein abundance are beneficial at low expression but could potentially incur a cost at high promoter activity. We thereby demonstrate that promoter activity in interaction with protein properties can dictate which coding mutations are beneficial, neutral or deleterious.

Main Text:

Introduction

Mutations drive evolution through multiple effects, from altering gene expression to modifying the stability, assembly and function of proteins, which ultimately affect cellular phenotypes and fitness. Therefore, measuring the fitness effects of mutations and identifying the underlying molecular mechanisms has been a long-standing goal in evolutionary biology (*1*, *2*). Due to epistasis, the effects of new mutations are dependent on prior mutations, which makes their impact on fitness difficult to predict. Epistasis can take place among genes or within genes (*3*), including between non-coding mutations that alter different features of a gene such as its regulation, including as promoter activity, and others that are coding and that alter protein activity or stability.

Regulatory and coding changes underlie phenotypic variation within and between species (*4*, *5*). Unfortunately, both types of mutations are rarely studied simultaneously when mapping the effects of mutations on fitness, apart from a few cases (*6–9*) revealing that epistasis between mutations affecting the expression of a gene and coding mutations (regulatory-by-coding) is likely frequent. For instance, the relative effects of coding mutations could be amplified at low expression level, as demonstrated in the study of a short amino acid segment of the heat shock protein Hsp90 (*9*). Conversely, mutations that are deleterious at low expression levels could become neutral at higher expression levels, or vice-versa. For example, some disease-associated mutations have a higher penetrance when linked to a regulatory variant that increases their expression (*10*). Thus, interactions between changes in gene expression and coding mutations could represent a vast underexplored constraint on protein evolution. Here, we examine how regulatory-by-coding epistasis could affect the evolutionary potential of a protein by performing a comprehensive measurement of the fitness effects of coding mutations at five levels of promoter activity in the vicinity of an expression level that is optimal for the wild-type (WT) enzyme.

Results and discussion

The fitness landscape of a protein complex depends on its promoter activity level

We investigated the extent to which promoter activity impacts on the fitness landscape of a protein by mimicking evolutionary changes by mutations affecting promoter activity using an inducible system. We used a metabolic enzyme that produces a metabolite that is conditionally essential for growth (Fig. 1A). The dihydrofolate reductase DfrB1 (formerly known as R67 DHFR) was first isolated as one of the factors causing trimethoprim (TMP) resistance in Enterobacteria, including in *Escherichia coli (11)*. Contrary to the monomeric FolA-type dihydrofolate reductase encoded in the *E. coli* main chromosome (ecDHFR) that is readily inhibited by TMP (ecDHFR K_i ~ 20 pM), DfrB1 is essentially insensitive to this antibiotic (DfrB1 K_i ~ 0.15 mM) (Fig. 1B) (*12*, *13*). The DfrB enzymes are obligate homotetramers that form a single, central active site requiring the distinct contribution from each of the four identical protomers (*12*, *14*). DfrB enzymes are not homologous to the ubiquitous FolA-type DHFR enzymes that ecDHFR is part of (*14–16*). While many antibiotic resistance proteins degrade or inactivate molecules, or expel them from the cell, DfrB1 is a metabolic enzyme that circumvents the inhibition of ecDHFR by duplicating the same essential enzymatic reaction: the reduction of dihydrofolate to tetrahydrofolate. The growth of *E. coli* can therefore be made dependent on DfrB1 activity through the inhibition of ecDHFR using TMP.

We used a plasmid whereby expression of *dfrB1* is driven by an arabinose-dependent promoter (Fig. 1B). By inducing expression in 0 to 0.4 % arabinose and measuring the extent of fitness recovery from ecDHFR inhibition (percentage of growth without TMP achieved with TMP), we found that WT DfrB1 promoter activity at 0.2 % arabinose completely complements the inhibition of ecDHFR. This induction level was defined as the optimal promoter activity level for WT DfrB1, and thus we will use this arabinose concentration as the reference for most comparisons (Fig. 1C, Fig. S1). Given the degree of complementation at each induction level, we refer to these levels as, starting from the highest,

above-optimal (0.4 % arabinose), optimal (0.2 % arabinose), near-optimal (0.05 % arabinose), suboptimal (0.025 % arabinose) and weak (0.01 % arabinose) for the WT enzyme. We measured the corresponding expression levels by fusing DfrB1 to GFP (Fig. 1D). *E. coli* populations displayed a unimodal expression level within this range of induction (median fluorescence in arbitrary units: 0 % arabinose: 0.42; 0.01 % arabinose: 0.74; 0.025 % arabinose: 1.11; 0.05 % arabinose: 2.55; 0.2 % arabinose: 16.07; 0.4 % arabinose: 18.85). DfrB1 therefore has a typical enzymatic fitness function as its fitness increases with promoter activity and then saturates when additional expression becomes superfluous (*17*).

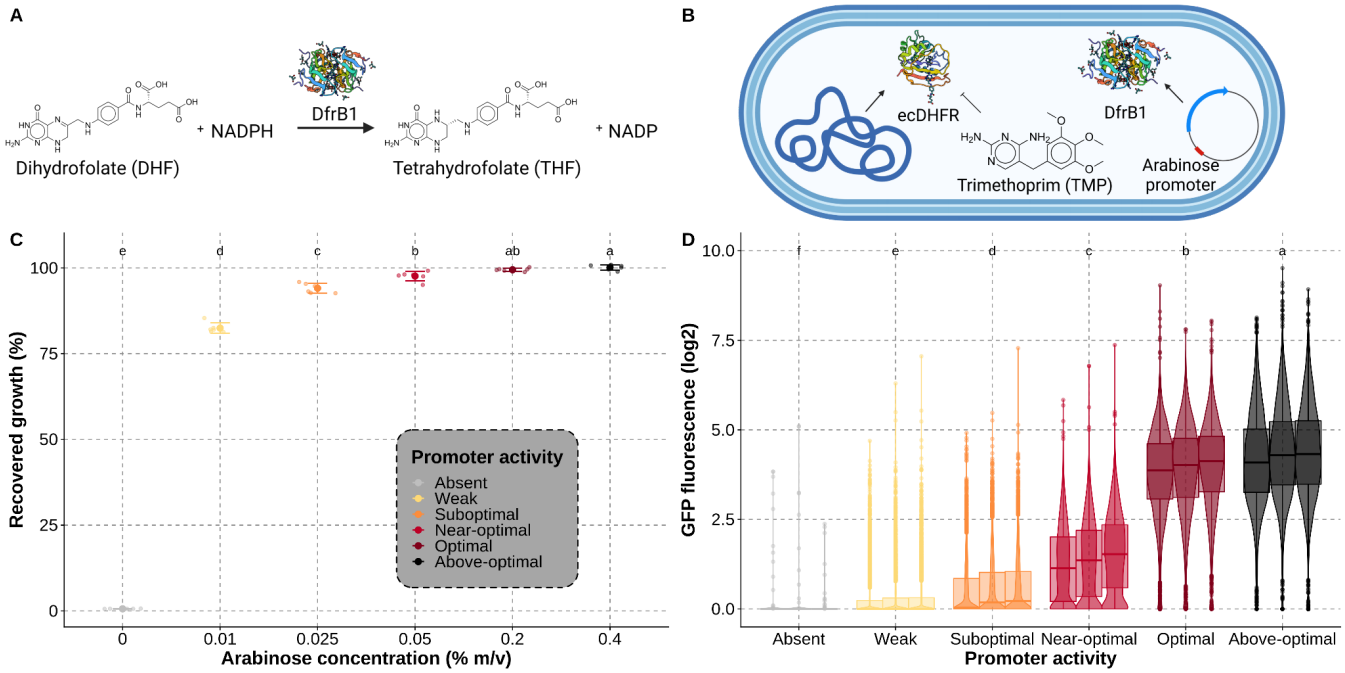


Fig. 1. Inducible DfrB1 expression connects promoter activity with fitness. (**A**) DfrB1 converts dihydrofolate to tetrahydrofolate, an essential metabolite. (B) The *E. coli* dihydrofolate reductase (ecDHFR) is inhibited by trimethoprim (TMP). DfrB1 is not homologous to ecDHFR and is insensitive to TMP, and, in our experiments, its transcription is modulated by an arabinose promoter. (C) Fitness function of DfrB1 promoter activity level. By inhibiting the endogenous ecDHFR with TMP (10 µg/mL), growth depends on the expression of DfrB1. Fitness is measured as a difference in growth (area under the curve) of the strain in medium with TMP relative to the same medium without TMP for a

given arabinose concentration. Dots indicate the mean of six replicates and error bars indicate 2.5 standard errors of the mean. Labels at the top indicate groups of recovered growth shown to be significantly different using an ANOVA and Tukey's HSD post-hoc test. (D) DfrB1-GFP expression as a function of arabinose concentration. Fluorescence of the WT DfrB1 is shown relative to the same *E. coli* strain but not expressing GFP. Three independent replicated cultures are shown side-by-side for each concentration. Labels at the top indicate groups of GFP fluorescence shown to be significantly different using an ANOVA and Tukey's HSD post-hoc test.

We generated a library of all possible single amino acid changes to the sequence of DfrB1 covering amino acid positions 2 to 78 (Fig. S2). Individual mutants were combined in a pool and transferred to medium with TMP at the five arabinose concentrations defined above (Fig. S3). The frequency of the mutants in the initial pool before selection and after 10 generations with TMP was measured by deep sequencing in a highly replicated fashion (Table S1, Table S2). We estimated a selection coefficient (s) per DNA variant at each induction level, and then redundant codons and replicates were aggregated (Fig. 2, Table S3). Selection coefficients are proportional to differences in growth per generation between cells with a specific DfrB1 mutation and cells with the WT DfrB1. Positive values of s indicate mutants that grow faster than the WT, negative values of s refer to mutants that grow slower than the WT, and a value of zero indicates the same growth rate as the WT (see Methods). Biological replicates were correlated across experiments, showing that the measurements are highly reproducible (Fig. S4). We confirmed the deleterious effects of several previously reported catalytically inefficient mutations (Fig. S5, Fig. S6). We also validated a number of isolated mutants ($n = 9$) through growth curves (Spearman $\rho = 0.87$, $p < 0.001$) (Fig. S7). These results confirm the validity of the pooled measurements to estimate fitness effects of variants of this enzyme.

The selection coefficients of individual mutants are broadly correlated among promoter activity levels, revealing that most amino acid substitutions have effects that are relatively similar across promoter activity levels. However, the magnitude of these correlations when the various induction levels are compared to the optimal level for the WT (0.2 % arabinose) decreases gradually as promoter activity level deviates from this reference (Fig. 2A-B, Fig. S8A). We observe a fitness effect that depends on promoter activity for many mutations: 598 out of 1458 substitutions (41%) (74/77 positions) measured at all five promoter activity levels showed statistically significant effects of promoter activity on selection coefficients (magnitude > 0.1, ANOVA, FDR adjusted $p < 0.05$, Fig. S8B, Table S4).

To identify broad patterns of changes of individual amino acid changes along the gradient of promoter activity, we grouped the vectors of selection coefficient by similarity. Clustering using k-means identified four typical patterns ($k = 4$ clusters). This number of clusters is the best compromise between parsimony and interpretability (Fig. 2C, Table S5, Fig. S9A-B) since from $k = 4$ onwards, adding more clusters does not result in a substantial decrease in the sum of squared errors. Representative centroids for each cluster are shown in Figure S9B. As one could expect, some substitutions do not have major effects on fitness (cluster 4, $n = 581$). At the other end of the spectrum, others are deleterious at all promoter activity levels (cluster 1, $n = 446$). Many variants appear to deviate from these two patterns (clusters 1 and 4), particularly between optimal (reference) and weak promoter activity levels (Spearman $\rho = 0.94$) (Fig. 2C). Mutants in cluster 3 ($n = 218$) are deleterious at weak and suboptimal promoter activity, but some become neutral or near neutral at the optimal promoter activity level for the WT enzyme. Mutants from cluster 2 ($n = 291$) are even more deleterious at low promoter activity levels but, even if they remain deleterious, they improve in fitness as promoter activity increases, even beyond optimal induction for the WT. This gain of fitness of many mutants at above-optimal promoter activity is also visible in the relationship shown in Fig. 2A. Mutations of intermediate deleterious effects at the optimal promoter activity level are worse as promoter activity is decreased and some are further buffered

or masked at above-optimal promoter activity. Finally, within cluster 4, some amino acid changes are advantageous at the lowest promoter activity level (see details below). Thus, our analyses confirm, both by statistical analysis (ANOVA for individual mutations, corrected for multiple hypothesis testing (Table S4, Fig. S8B)) and by machine learning (k-means clustering (Table S5, Fig. S9)), widespread epistasis between promoter activity level and coding mutations, with promoter activity levels having different impacts on the selection coefficient of individual mutations.

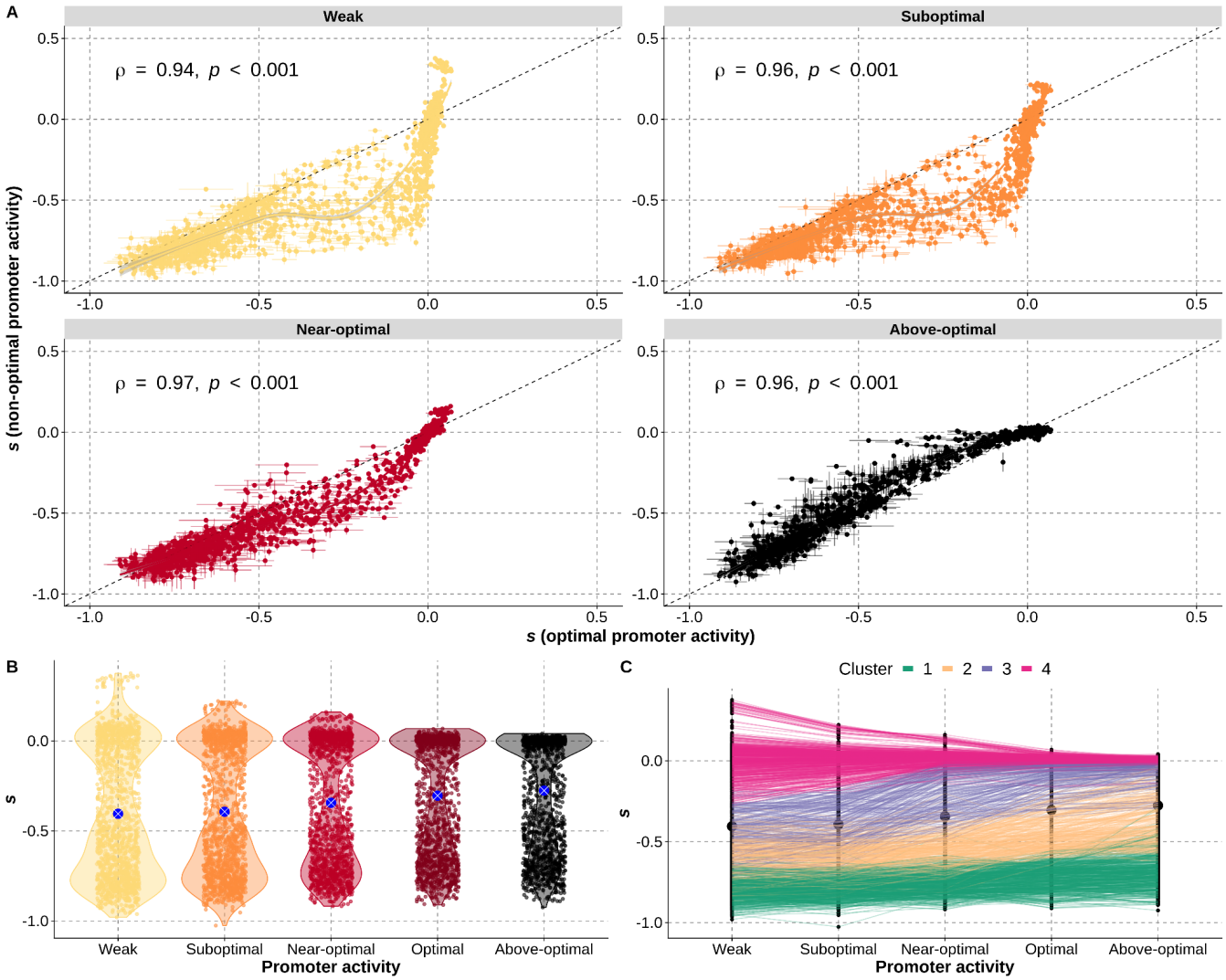


Fig. 2. Selection coefficients (*s*) of amino acid changes depend on promoter activity. (**A**) Selection coefficient (*s*) of individual amino acid mutants at non-optimal levels of promoter activity compared to the optimal level for the WT, which is used as a reference point for comparisons. Spearman correlation coefficients are shown. Lines indicate LOESS regressions. Horizontal error bars indicate the standard

error of the mean observed across biological replicates at optimal promoter activity level for the WT. Vertical error bars indicate the standard error of the mean observed across biological replicates at non-optimal promoter activity levels. (B) Distributions of *s* for individual mutants across promoter activity levels show reduced spread and increasing average of individual values. The larger blue dots with white crosses indicate the means of the distributions. (C) Clusters showing patterns of changes of *s* for individual mutants across promoter activity levels. Colored lines indicate mutants from the different clusters (k-means, k = 4). Number of mutants in each cluster: 1) 446, 2) 291, 3) 218, 4) 581.

We also examined whether mutants changed fitness ranking among promoter activity levels. A change in ranking between two mutations implies that a first mutation would outcompete a second one at a given promoter activity level, but that the second mutation would outcompete the first at a different promoter activity level. An ANOVA on the ranks revealed a significant interaction between genotype and promoter activity level ($F = 12.15$, $p < 2.2e-16$), although these effects are much smaller than those associated with mutations alone ($F = 666.10$, $p < 2.2e-16$) (Table S6). To validate this change in ranking, we tested a few mutants ($n = 8$) changing ranks in the bulk assay in individual growth curve assays. As previously, mutants with higher fitness at optimal promoter activity for WT were clearly reproduced. Mutants that were less deleterious at lower promoter activity and that changed rank in the bulk competition experiment were not as consistent with the validation cultures. For example, in the bulk competition experiment, E75R was estimated to be the second most deleterious substitution of the set considered at weak promoter activity but become less deleterious at optimal expression than mutations at positions 33 and 68 (K33E, K33L, K33M, I68M). Such changes in rank would be visualized as crossing lines on Fig. S10. E75 did indeed become less deleterious than K33L and K33M at optimal promoter activity level in the validation cultures, but we did not observe the expected change in rank of E75 with respect to K33E and I68M (Fig. S10). In particular, K33E had a higher fitness at weak promoter activity in our bulk competition experiment but a higher fitness at optimal promoter activity in

the validation cultures. This means that it is more challenging to measure precise selection coefficients for highly deleterious amino acid substitutions in bulk competition. Although changes of rank may occur, it is not the major type of expression-by-coding epistasis observed for DfrB1. Indeed, the fact that deleterious mutations are strongly underrepresented at the final timepoint in bulk competition experiments increases experimental noise. Nevertheless, the differences in fitness between promoter activity levels in the individual versus bulk competition assays were strongly correlated (Spearman $\rho = 0.82$, $p = 0.011$) (Fig. S10).

The protein structure determines which amino acid changes have effects that depend on promoter activity level

Individual amino acid mutations can alter many protein properties (*18*) and thus decrease the amount of functional protein complex formed, which can result from a reduction in the stability of the protein or of its tertiary assembly. The solubility of the protein can also be altered by individual mutations. For enzymes, individual amino acid changes can alter catalytic efficiency. To examine the potential underlying causes of the patterns revealed in Fig. 2, we mapped selection coefficients of all amino acid substitutions to individual protein positions (Fig. 3). Changes at conserved positions tend to be more deleterious than in non-conserved regions when considering diversity across homologs (Entropy, Fig. 3 - top; Spearman correlations between entropy and selection coefficients range between $\rho = 0.82$ and $\rho = 0.83$ across promoter activity levels, all with $p < 4e\text{-}20$) or using predictions from patterns of molecular evolution (*19*) (Fig. S11). Selection coefficients as measured in the laboratory therefore correlate with the constraints acting on the various positions of the protein in nature, as estimated by their degree of evolutionary conservation.

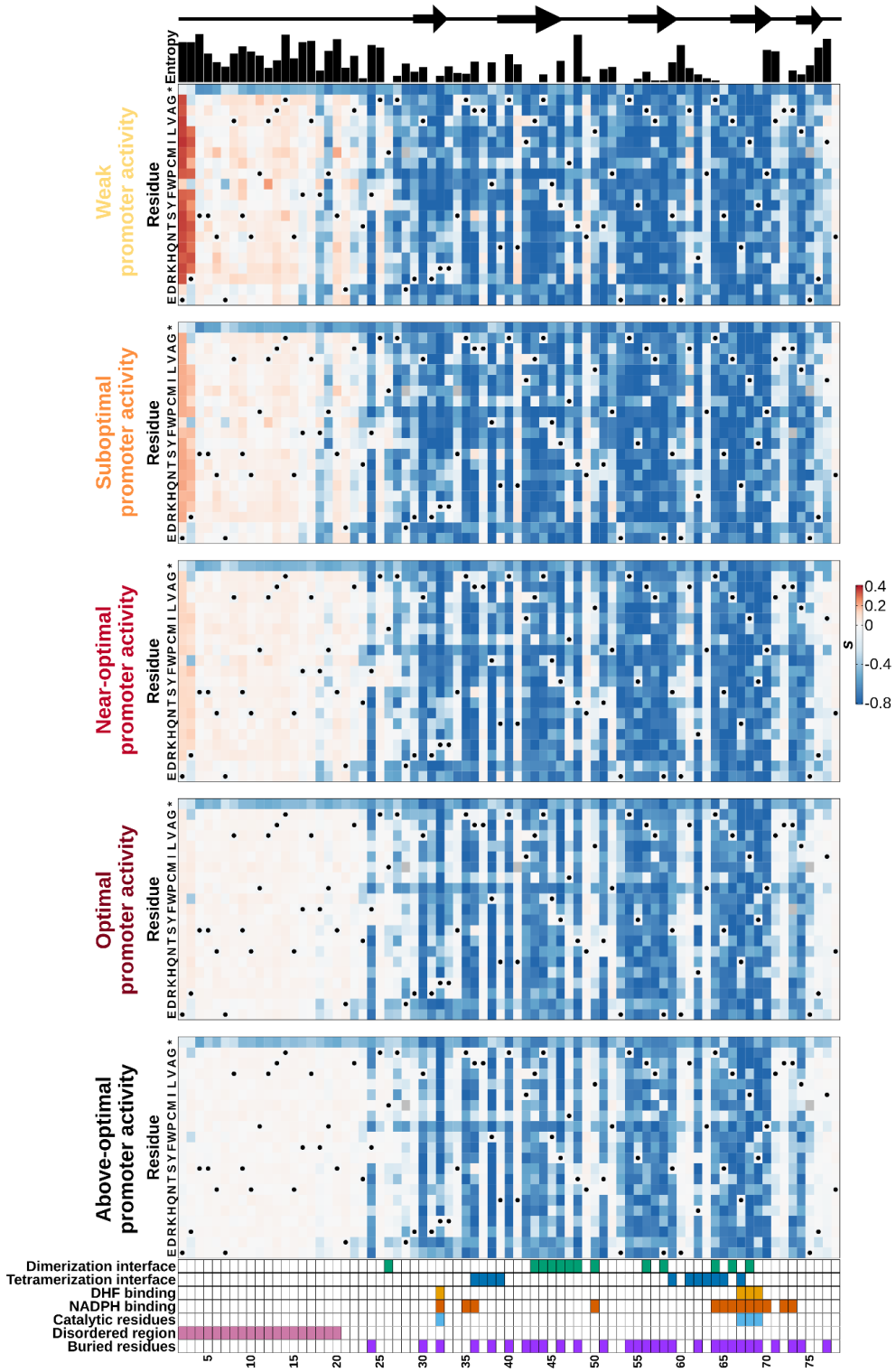


Fig. 3. The DfrB1 fitness landscape changes with promoter activity level in a very heterogeneous way along the protein sequence. Selection coefficient (*s*) of individual amino acid substitutions (rows) at different positions (columns) at five promoter activity levels. Annotations above the first panel

indicate the protein's local secondary structure (arrows indicate beta strands) and sequence diversity among homologs (Shannon entropy). The annotations at the bottom of the last panel highlight residues that participate in the interfaces of the DfrB1 homotetramer, that come into contact with either the substrate (DHF) or the cofactor (NADPH), key catalytic residues, the disordered region, and buried sites. Interfaces are labeled as dimerization and tetramerization interfaces following (*22*). Residues marked with a dot on the heatmaps are wild-type residues (WT codons and their synonymous ones).

The fitness effects of amino acid changes are heterogenous along the length of the sequence and depend on the structural property of each segment. For instance, substitutions in the disordered region (positions 2-20) (*15*, *20*) tend to have a higher selection coefficient (less negative or more positive) than the rest of the protein ($p < 2.2e-16$ for all promoter activities, Wilcoxon test). Results are consistent with the higher entropy of the disordered region ($p = 4.2e-06$, Wilcoxon test) and with results showing that the sequence essential for DfrB1 function begins between residues 16-26 (*21*). Importantly, mutants from specific protein functional sites marked at the bottom of Fig. 3 tend to be enriched in different clusters (Fig. 2C) of selection coefficients across promoter activity levels (Fig. S9C). For example, substitutions in the disordered region are enriched in cluster 4 (mostly neutral) and depleted in cluster 1 (always deleterious) (chi-squared test, $p = 6.2e-116$). On the other hand, changes in the catalytic sites are particularly enriched in cluster 1, while substitutions at the dimerization interface primarily belong to clusters 1 and 2 (slight recovery with higher promoter activity) (chi-squared test, $p = 2.3e-18$). This enrichment of negative effects at catalytic sites is consistent with the strongly negative selection coefficients of most known catalytic mutants of DfrB1 (Fig. S5). These observations show that the extent to which promoter activity level can modify the fitness effects of mutations can be explained by the various structural features of the enzyme (see more details below).

Comparison of the landscapes at the five promoter activity levels (Fig. 3) confirms a dampening of selection coefficient from weak to above-optimal promoter activity, showing that extreme fitness effects, positive as well as negative, become rarer as promoter activity level approaches the optimum for the WT. This again illustrates that the effects of different amino acid substitutions tend to be buffered at promoter activity levels near the optimum for the WT enzyme. A change of promoter activity level changes which positions of the enzyme become visible to selection as the dampening of selection coefficients is site-dependent.

To quantify the effect of promoter activity level on selection coefficients, we calculated Δs as the difference between the selection coefficients of an amino acid substitution at a given promoter activity level and at the optimal level for the WT enzyme that we use as a reference ($\Delta s_{non\text{-}opt} = s_{non\text{-}opt} - s_{opt}$, with $s_{non\text{-}opt}$ referring to either of s_{weak}, s_{subopt}, $s_{near\text{-}opt}$, or $s_{above\text{-}opt}$). Heatmaps similar to Fig. 3 but with the $\Delta s_{non\text{-}opt}$ scores (Fig. S12A) show that: 1) a majority of positions with strongly deleterious mutations are deleterious no matter the promoter activity level ($\Delta s_{non\text{-}opt}$ around 0), 2) the deleterious effect of some mutations is dampened at the reference promoter activity level ($\Delta s_{non\text{-}opt} < 0$), and 3) some mutations are beneficial specifically at low promoter activity ($\Delta s_{non\text{-}opt} > 0$). The dampening of mutational effects by gene expression is dose-dependent since the magnitude of $\Delta s_{non\text{-}opt}$ becomes smaller as promoter activity level approaches the optimum, so mutational effects observed at weak promoter activity would start to be dampened even at suboptimal promoter activity ($p \leq 1.9e\text{-}09$ for all comparisons, Wilcoxon test for differences in means of paired samples, Fig. S12B). We chose to focus on Δs_{weak} in the following sections because it represented the most dramatic decrease in promoter activity with respect to the optimum.

Low promoter activity makes some coding mutations beneficial

The disordered N-terminal region is distinct from the rest of the protein as it contains beneficial amino acid changes whose effects disappear at the reference promoter activity level ($\Delta s_{weak} > 0$), particularly at positions 2 and 3. The E2R and E2V mutations for instance have extremely high values (Fig. 4A). This suggests that the enzyme may be more catalytically efficient or at higher abundance than the WT when promoter activity is low. To distinguish these two possibilities, we assayed enzyme activity in cellular extracts for two mutants at position 2 that had positive selection coefficients, particularly at the lowest promoter activity level (s = 0.340 for E2R and s = 0.363 for E2V). We confirmed an increase in bulk enzymatic activity for both E2R (p = 5.7e-3, t-test) and E2V (p = 0.05, t-test) relative to the WT (Fig. 4B). Interestingly, although only marginally significant (p = 0.09), E2R had higher activity than E2V, which corresponds to the slightly higher fitness observed for E2R in individual growth assays (Fig. S7A).

We next examined if the higher observed activity of the E2R and E2V mutants could be caused by higher protein abundance. Nucleotide content and their corresponding amino acids in the first few codons of a gene are known to modulate protein expression in *E. coli* (*23*), which makes this region a target for beneficial mutations when transcription levels are low. To test the effect of amino acid changes on protein abundance, we generated GFP fusions with: 1) the entire DfrB1 WT sequence with and without 2) the most beneficial mutation (E2R), 3) the WT disordered N-terminal region (amino acids 1-25), and 4) the disordered region with the E2R mutation (Fig. S13). When the E2R substitution was introduced, we observed higher protein abundance, even when only fusing the disordered fragment containing E2R to GFP, reaching more than a 10-fold expression change relative to the WT sequence (ANOVA p < 2.2e-16, Tukey HSD post-hoc test p < 1.3e-6) (Fig. 4C, Table S7). At the lowest promoter activity level (0.01 % arabinose), the E2R mutant fused to GFP has an abundance that compares or that is slightly higher than the WT enzyme at its optimal induction level. Such a beneficial mutation at low promoter activity therefore acts by increasing protein abundance, most likely by increasing translation

rate. The benefit extends to even lower promoter activity as growth recovery is maximal for this mutant at even lower induction level (Fig. 4D). This result demonstrates that the promoter activity and coding sequence could co-evolve to tune protein abundance. Thus, the optimal abundance of the protein could be achieved with lower promoter activity for the E2R and E2V mutants than for the WT. Conversely, when promoter activity approaches the optimum for the WT, the E2R and E2V mutations could lead to a decrease in fitness due to cost of producing an excess of the enzyme, or other potential negative impacts. We did observe a slight significant fitness reduction for the E2R mutant at high induction levels (Fig. 4D, Fig. S14). This would represent one of the few cases of changes in promoter activity that lead to a change in fitness ranking among mutants.

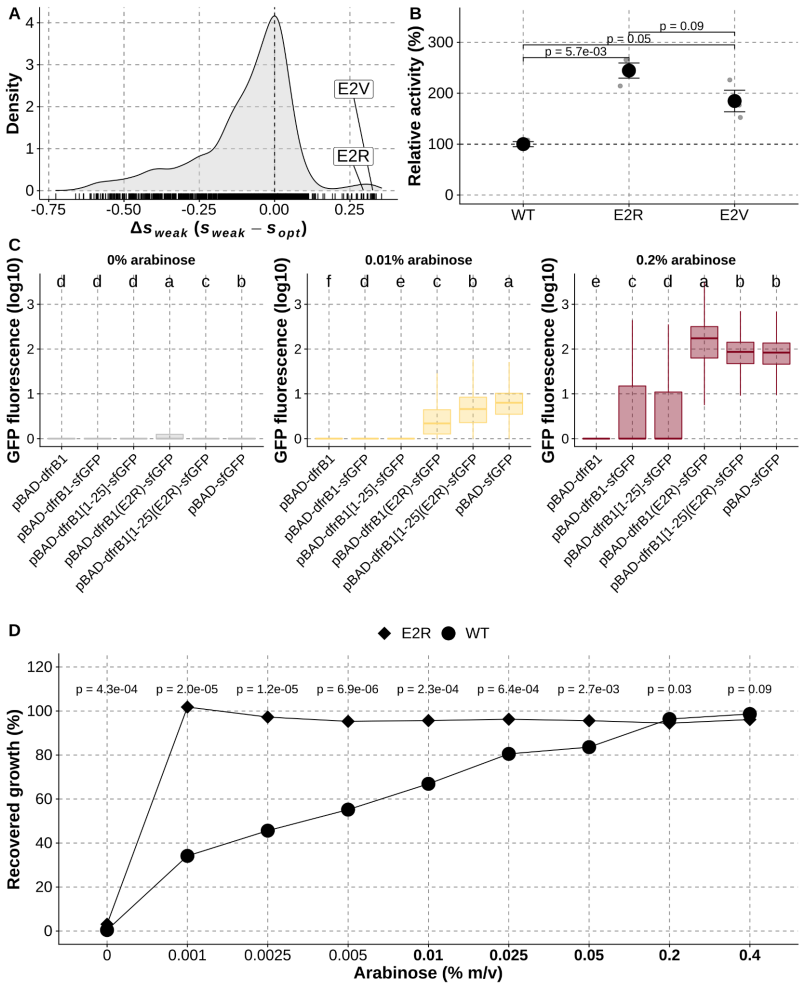


Figure 4. Mutations that are beneficial at low promoter activity lead to increases in enzyme activity through increased protein abundance. (A) Distribution of differences in fitness effects

observed for the same mutation at weak and optimal promoter activity. E2R and E2V have strongly beneficial selection coefficients at weak promoter activity, leading to positive Δs_{weak}. (B) Lysates from cells expressing the E2R and E2V mutants show a higher *in vitro* bulk activity compared to WT DfrB1. Error bars indicate the standard error of the mean of three replicates. P-values were calculated using t-tests. (C) Fused DfrB1-GFP constructs have a higher protein abundance when they harbor the E2R mutation than when harboring the WT truncated DfrB1 sequence. Data show the distributions of fluorescence intensity for 5000 cells. Labels at the top correspond to groups that show significant differences based on an ANOVA and Tukey's HSD post-hoc test. (D) The E2R mutation confers a higher growth recovery even at promoter activity levels far below the ones used for the bulk competition experiment. Dots indicate the average of three replicates. P-values were calculated from t-tests comparing growth recovery of WT DfrB1 and E2R mutant. Concentrations of arabinose in bold indicate induction levels used in the bulk competition experiment.

Masking of fitness effects is encoded in the protein structure

Unlike the individual selection coefficients at each level of promoter activity, the average Δs_{weak} does not correlate with the entropy measured for each position in the alignment (Fig. 5A). Instead, the positions with the strongest masking effects are those on the outside of the protein and those with no clear function (Fig. 5B - mean Δs_{weak}). However, strongly negative values of Δs_{weak} are frequent at buried positions (Fig. 5B - minimum Δs_{weak}), and the most strongly positive values are found in the disordered region (Fig. 5B - maximum Δs_{weak}). As a result, promoter activity-dependent effects are observed throughout the protein structure.

The distributions of Δs_{weak} for each kind of protein site show a more complete picture. Mutations that could destabilize the protein complex, such as those in buried sites and at the protein interfaces (Fig. 5C), seem to be buffered by increased promoter activity. An increase in expression level could

compensate for lower stability following the law of mass action by augmenting the concentration of functional complexes. Even the sites in the disordered region with negative Δs_{weak} (F18 and P19) (Fig. 5A, Fig. S15) could be explained by interaction interfaces. AlphaFold2 (25, 26) predicts residues F18 and P19 to be in contact with W45 (Fig. S15A-B), which is known to interact with the disordered N-terminal region (27). These analyses show the sharp contrast between mutants according to their structural and enzymatic roles: those involved in catalysis having Δs_{weak} near 0, those in the disordered region often having positive Δs_{weak} and those at interaction interfaces with more negative Δs_{weak} (Fig. 5C). Indeed, these contrasts are reflected in the variability observed in extant sequences, with the disordered region having higher entropy values than the interfaces and catalytic sites (Fig. 5A).

An important consideration is that $\Delta s_{weak} = 0$ could indicate both mutations whose effects are not sensitive to promoter activity, but also those that are completely neutral. For instance, the Q67C mutant has a catalytic efficiency reduced by 100-fold with respect to the WT DfrB1 (24) and has a Δs_{weak} near 0 ($\Delta s_{weak} = -0.008$), but the same is true for many mutations in the disordered region. We therefore mapped the effects of mutations in a two dimensional space where Δs_{weak} values are plotted against s at optimal promoter activity for the WT (Fig. 5D). This representation confirms the striking difference between promoter activity-independent effects of substitutions in catalytic residues and expression-modulated effects at interaction interfaces (ANOVA $p < 2.2e-16$, Tukey HSD post-hoc test $p < 0.037$) (Fig. 5C and E, Table S8). These differences remain significant when separating the data into residues that are exclusively annotated as catalytic, those that are exclusively annotates as interfaces, and those that are annotated as both catalytic and interface residues (ANOVA $p = 0.036$, Tukey HSD post-hoc test $p = 0.029$). Mutants with poor catalytic activity may not benefit much from change in promoter activity in this range. Of the mutants with known k_{cat}/K_M, only the one with a limited (36.5 %) reduction in activity S65A (S65A $k_{cat}/K_M = 0.275$; WT $k_{cat}/K_M = 0.433$) appears to be significantly improved by increasing promoter activity (S65A $\Delta s_{weak} = -0.21$; effect of promoter activity level on fitness, FDR adjusted

p-value = 5.59e-06, Table S4). All others, with k_{cat}/K_M under 3 % of the WT, have Δs_{weak} near 0 (Q67C Δs_{weak} = -0.018, I68L Δs_{weak} = -0.025, I68M Δs_{weak} = 0.014, Y69F Δs_{weak} = 0.024, Y69H Δs_{weak} = -0.021). Mutations that reduce enzyme activity down to a few percent of that of the WT are therefore unconditionally deleterious in this range of promoter activity.

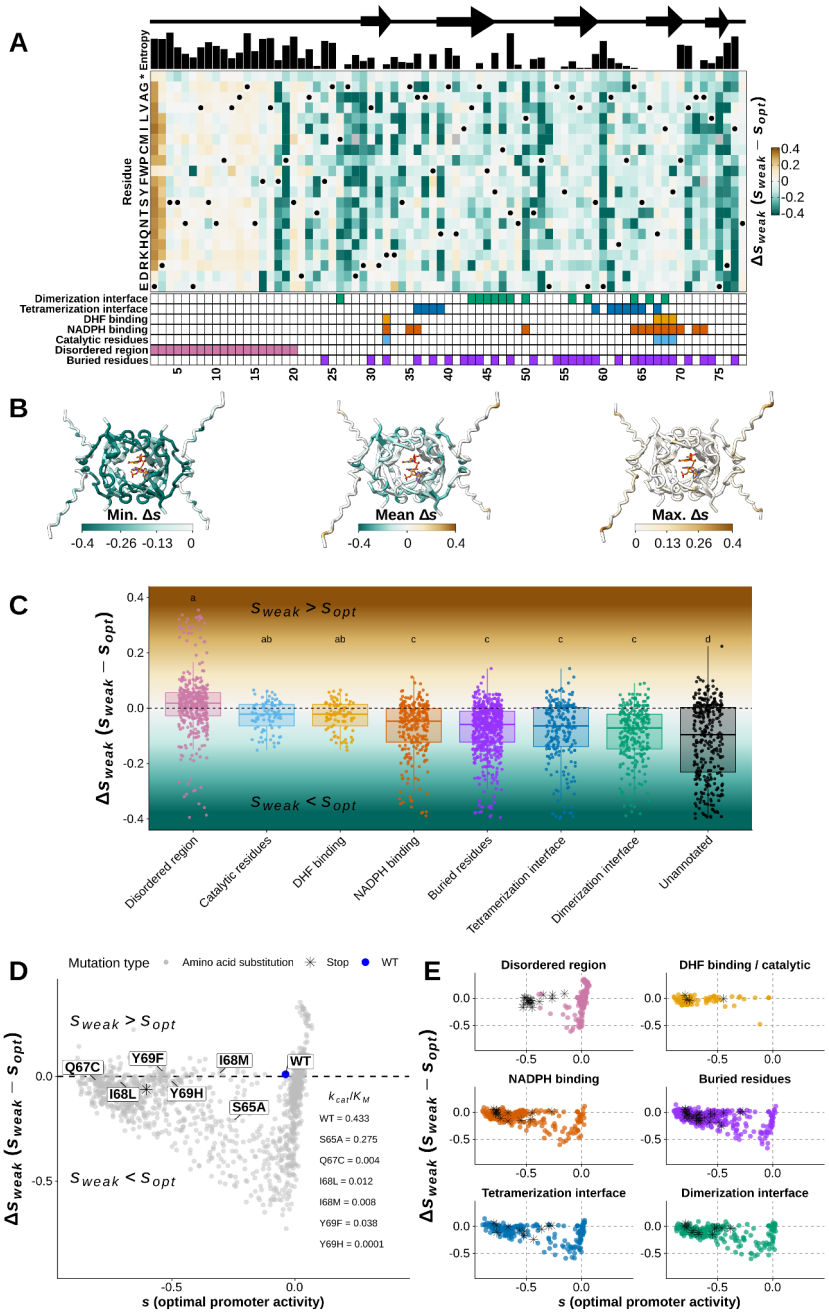


Fig. 5. Optimal promoter activity for the WT enzyme buffers mutational effects to varying degrees depending on the structural and catalytic sites. (**A**) Δs for weak versus optimal promoter activity

level (Δs_{weak}). A negative Δs_{weak} value indicates mutations that are more advantageous at optimal promoter activity for the WT. A positive Δs_{weak} value indicates mutations that are more advantageous at weak promoter activity relative to the WT. Annotations at the top (Shannon entropy) and bottom are the same as in Fig. 3. (B) Minimum, mean and maximum Δs_{weak} values mapped on DfrB1 structure. The structures were generated with AlphaFold2 to include a representation of the disordered region, which is not present in PDB: 2RK1. The molecules in the active site represent the substrate (DHF) and cofactor (NADPH) for the enzymatic reaction using the coordinates from PDB: 2RK1. For all other structural analyses, the coordinates from PDB: 2RK1 were used. (C) Δs_{weak} values vary with structural properties. The shaded background corresponds to the color scale from Fig. 5A. Letters above boxplots indicate groups with statistically significant differences as shown with an ANOVA and Tukey's HSD post-hoc test (Table S8). Positions belonging to more than one site are included in the individual distributions for all of those sites. (D) Overall landscape of selection coefficients at optimal promoter activity for the WT (s_{opt}, x-axis) and Δs_{weak} (y-axis). Mutants close to the dashed horizontal line have fitness effects that are independent from promoter activity. Labels indicate mutants with experimentally determined catalytic efficiency (k_{cat}/K_M) from Strader et al. (*24*), also shown in Fig. S5, S6B. (E) Functional sites of the protein occupy distinct regions of the selection/buffering landscape. Stop codons are labeled as black asterisks.

Slightly destabilizing mutants are masked at optimal promoter activity for the WT

The previous analyses revealed that buried sites and binding interfaces often show mutations with effects that can be dampened by increasing promoter activity, suggesting that one of the causes of expression-dependent effects is protein destabilization. Protein destabilization would reduce the amount of protein complex formed for a given level of promoter activity by two major mechanisms: destabilizing individual protein subunits or by reducing the binding affinity at the interfaces of the homotetramer (Fig. 6A). These factors can contribute to reducing active enzyme abundance the same

way changes in promoter activity or translation do. To further examine the contribution of protein destabilization to the observed Δs_{weak}, we computationally estimated the effects of amino acid changes on the protein and complex stability. Overall, we find a weak but significant correlation between $\Delta\Delta G$ of subunit stability and Δs_{weak} (Spearman ρ = 0.10, p-value = 4.3e-4). We mapped destabilizing effects in the landscape of fitness effects at optimum promoter activity (s_{opt}) versus changes in fitness effects caused by a weak promoter activity (Δs_{weak}) (Fig. 6B). This analysis shows that highly destabilizing mutations, especially those affecting subunit stability, cluster in the same region of the landscape in Fig. 6B as those with poor catalytic activity in Fig. 5D and stop codons in Fig. 5E, corresponding to mutations with limited or no masking capacity at optimal promoter activity for the WT.

Since the same mutation can have an effect on both protein stability and binding affinity at either of the interfaces, we looked for mutations affecting only one of these parameters at a time. For example, to focus on protein stability, we restricted further analyses to mutations causing changes between -0.5 and 0.5 kcal/mol on binding affinity at the interfaces. We did the same for mutations affecting the dimer and tetramer-forming interfaces. We observed that higher promoter activity levels tend to improve fitness for all mutations but to a very limited extent for those with $\Delta\Delta G$ greater than 2 kcal/mol, both on protein stability and binding affinity at either interface, which largely behave like stop-codon mutants (Fig. 6C, Fig. S9C). Indeed, substitutions with highly destabilizing effects on subunit stability (p = 1.6e-47) and on binding affinity at the dimer (p = 8.3e-16) and tetramer (p = 8.6e-4) interfaces tend to be mostly in clusters 1 (always deleterious) and 2 (limited buffering by WT expression) as seen from the k-means analysis (p-values were estimated using chi-squared tests and adjusted using the Benjamini-Hochberg correction with FDR < 0.05, Fig. S9C). In contrast, substitutions with weakly destabilizing effects on subunit stability ($\Delta\Delta G_{stab}$ lower than 2 kcal/mol) belong mostly to clusters 3 (strong buffering by WT expression) and 4 (mostly neutral) (Fig. S9C). These observations also help explain the differences in the distribution of Δs_{weak} between buried and unannotated sites (Fig. 5C) because mutations in buried

sites tend to be more destabilizing than mutations in exposed sites. To confirm the predominant role of protein destabilization, we trained a random forest regressor (*25*) (see methods) on Δs_{weak} using the predicted biophysical effects of amino acid changes, structural features, changes in amino acid properties based on the differences in 57 indexes from ProtScale (*26*), and the propensity of amino acids to be found in protein-protein interaction interfaces (*27*) (Table S9). While the global performance was modest (R^2 (test set) = 0.44, R^2 (5-fold cross-validation) = 0.27 ± 0.031 (standard error of the mean)), it identified relative solvent accessibility, changes in hydrophobicity and effects on subunit stability as the top features contributing to the predictions (Fig. S16). Finally, an ANOVA modeling selection coefficients based on expression level and bins of destabilizing effects found a statistically significant interaction between promoter activity levels and effects on subunit stability (p = 0.017) (Table S10). As a result from these two analyses, we conclude that the relation between destabilizing effects of mutations and promoter activity is likely nonlinear, which is illustrated in Fig. 6C. As protein destabilization increases, the improving effect of increased promoter activity on fitness decreases.

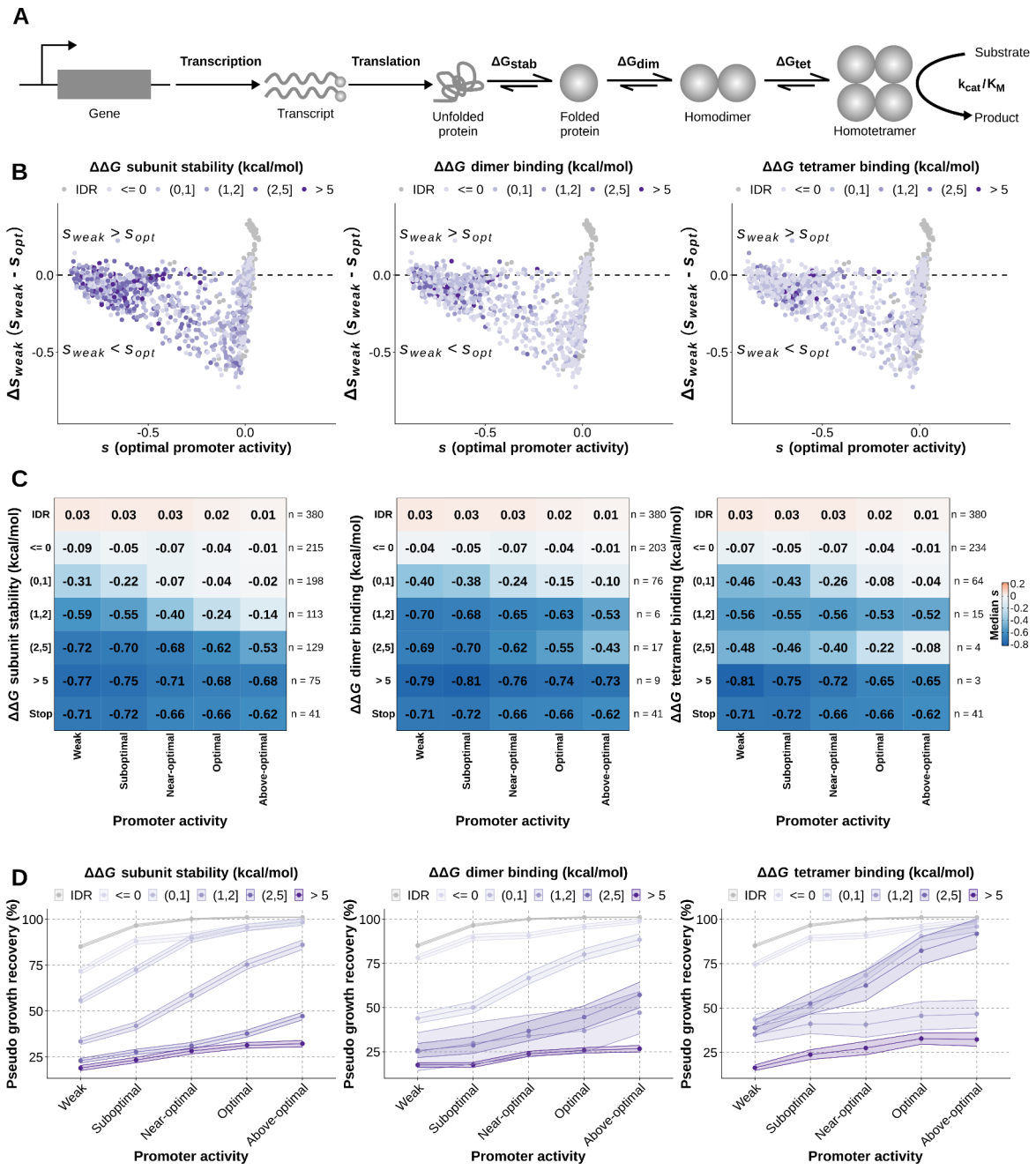


Fig 6. Increased promoter activity masks slightly destabilizing mutations but not highly destabilizing ones. (**A**) Mutations can affect many levels of protein regulation and activity, from transcription, translation to enzyme activity through protein stability or catalytic activity. (B) Landscape of selection coefficients at optimal promoter activity for the WT (x-axis) and Δs_{weak} (y-axis). Substitutions are represented as dots and colored according to bins of the magnitude of their destabilizing effects on subunit stability and binding affinity at the interfaces. Substitutions in the

disordered region (positions 2-20) are also included for reference as a category of their own because they are not present in the PDB structure (2RK1) and hence their destabilizing effects would not be reliably estimated. Highly destabilizing mutations cluster on the left where mutations are deleterious (negative s) and the deleteriousness tends to be independent of promoter activity level (Δs_{weak} around 0). (C) Heatmaps showing medians of measured selection coefficients for mutants with different effects on protein stability or binding affinity at the interfaces. Stop codons from the functional protein core (residues 30 - 70) are shown for comparison. Numbers of mutations considered for each bin of effects on protein stability or binding affinity at the interface are listed on the right. Increases in promoter activity lead to an increase in the median fitness of mutations with slightly destabilizing effects, but not as much for those with highly destabilizing effects. (D) Pseudo growth recovery functions for mutations from the different bins of effects on protein stability and binding affinity. Dots indicate the median and shaded ribbons indicate the standard error of the mean.

The previous section considered that protein destabilization would act by reducing protein abundance or catalytic activity, which would in turn reduce fitness (Fig. 6A). Fitness reduction could also come from an unstable protein through spurious interactions with other proteins (*28*), leading to toxic effects that would increase with promoter activity. We confirmed that expression of WT DfrB1 was not toxic in itself at any of the tested promoter activity levels in the absence of TMP (Fig. S17). Although it is difficult to infer protein misfolding from the $\Delta\Delta G$ predictions, in principle, misfolded proteins that lead to toxicity would reduce fitness even in the absence of TMP. We therefore repeated the bulk competition experiment of the mutant library, inducing expression in 0 to 0.4 % arabinose, but this time removing selection for DfrB1 activity by omitting TMP. We observed lower signal-to-noise ratio, with a lower correlation between biological replicates (Fig. S18) than when TMP selected for DfrB1 activity, consistent with very little fitness differences among mutants. The distributions of selection coefficients are tightly clustered around 0, confirming the lack of detectable fitness effects of most mutations when

selection is removed (Fig. S19, S20). In addition, these selection coefficients do not correlate with patterns of molecular evolution (Fig. S21) and do not reflect any underlying effects on subunit stability or binding affinity (Fig. S22). As a result, most of the fitness effects measured in the experiment with TMP above are therefore unlikely to come from toxic misfolding and gains of interactions, but to a lowered activity or abundance of the functional tetrameric enzyme.

Our results show that the stability of the protein and of its complex would create classes of promoter activity-fitness functions different from the one observed for the WT enzyme (Fig. 1C). Because the estimated selection coefficients correlate well with growth recovery measured in growth curve assays (Fig. S7), we could infer fitness functions for individual mutants with different estimated stability classes. Using the WT promoter activity-fitness function (Fig. 1C) as a reference and the selection coefficients of individual mutations (Table S3), we derived a pseudo growth recovery metric (see Methods) to visualize the interaction between promoter activity level and mutational effects on protein stability and binding affinity on fitness (Fig. 6D). Growth recovery is higher as promoter activity increases, and this improvement tends to be higher for mutations with smaller effects on either protein stability or binding interfaces. Conversely, highly destabilizing mutations have fitness functions that are nearly flat, suggesting a limit to the buffering capacity of the optimal promoter activity.

Conclusion

By measuring how promoter activity impacts the fitness effects of amino acid substitutions on a small protein complex, we find that the selection coefficients of individual amino acid substitutions are highly conditioned by the activity of the promoter, revealing rampant regulatory-by-coding epistasis. Some coding mutations become strongly advantageous at low promoter activities, some mutations are strongly deleterious at any promoter activity and some have deleterious effects that are dampened by higher promoter activity. Coding mutations that are advantageous specifically at low promoter activity increase

protein abundance. Mutations that are unconditionally deleterious tend to have a strong negative impact on the catalytic activity of the enzyme or extreme effects on its stability.

Mutations with deleterious effects that are modulated by promoter activity appear to be slightly destabilizing for the protein itself or its assembly into dimers and a tetramer. Promoter activity could therefore potentially evolve to increase the abundance of slightly underperforming enzymes. Similarly, a recent study showed that the effects on fitness of slightly destabilizing mutations in a metabolic enzyme can be compensated by a higher availability of the substrate (*29*). As for the highly destabilizing mutations and those affecting the catalytic sites of the enzyme, they have selection coefficients that are less likely to be buffered by increasing promoter activity within the range examined. We cannot exclude that more destabilizing mutations could also be buffered at higher promoter activity. The range of expression that we considered is from 7 % to 110 % of optimal expression for the WT enzyme (Fig. 1D), to mimic the range observed for naturally occurring substitutions in promoter sequences (*30*), which likely is representative of what is available through single-step mutations. If mutations increasing promoter activity to a much higher level are not accessible, such coding mutations would never have a chance to be maintained, even more so since the cost of overexpression could be a significant burden for some proteins.

Our inferences come with shortcomings. For instance, for most mutants of the catalytic sites we have not determined experimentally how these changes affect catalytic efficiencies. However, by curating the literature, we identified many mutants whose catalytic efficiencies have been determined and those indeed remain strongly deleterious at all the levels of promoter activity we examined (Fig. S5). Similarly, computational estimations of mutational effects generally have a good agreement with experimental values but are not always as accurate (*31*, *32*). FoldX assumes a rigid backbone, which restricts the study of misfolding induced by mutations (*31*). Nevertheless, our FoldX predictions agree

well with the fitness estimations from our bulk competition experiment and reflect the biological effects of losses of bulk enzymatic activity.

Many biological parameters other than promoter activity can influence how much functional enzyme is produced. The total amount of protein in the cell therefore depends on the interplay and co-evolution of promoter activity, translation rate, and protein stability. For instance, many combinations of transcription and translation rates can lead to the same amount of protein (*33*). Coding mutations that increase protein abundance, for instance through codon usage (*23*), can therefore be beneficial at low promoter activity, as we observed here. Even if they produce similar steady-state amounts of proteins, combination of transcription and translation rates and protein stability are not evolutionarily equivalent in the short term, due to selection on other features such as expression noise (*33*). In the long term, other combinations of parameters can enhance evolvability, as a more lowly expressed but more stable protein was recently shown to have access to different adaptive paths (*34*). Being able to measure how mutations affecting transcriptional, post-transcriptional, translational and post-translational regulation interact to shape fitness will help better understand how they evolve in concert.

Our results show that changes in gene regulation such as alteration of promoter activity by mutation, could dictate which mutations will be beneficial or deleterious for a coding sequence, and vice-versa. Over the long term, promoter activity and protein sequences will thus co-evolve. This co-evolution could be disrupted when major changes occur in the environment and in the regulation of the expression level of a gene, such as in the event of gene duplication, translocation or horizontal gene transfer. These changes could be advantageous but not necessarily optimal, for instance in the case of drug resistance enzymes such as the one presented here. Reaching optimality could be achieved by a change in regulation of transcription, translation or protein stability. Adaptive protein evolution could therefore occur because of sub-optimal expression level or vice-versa. While we focused our analyses on protein

stability and the formation of a protein complex, other protein properties such as solubility could play a role in determining the optimal promoter activity for a particular variant. Overall, these results call for the joint consideration of coding and regulatory evolution in the study of protein evolution.

Although focusing on one particular protein, our results may shed light on a long-standing observation in molecular evolution. Highly expressed proteins evolve slowly (*35*, *36*) and this is often interpreted as higher selective constraints acting on highly expressed proteins, for instance to prevent misfolding and potentially misinteraction, which would be costly. We found potential misfolding to have no detectable effects for this protein, which could be the case for many other proteins as well. For such proteins, if high promoter activity buffers the effect of destabilizing mutations, one would wrongly expect highly expressed proteins to evolve faster. However, since highly expressed proteins are costly to produce (*37*), their expression level may actually be below the optimal level that would be observed in the absence of metabolic cost. As a result, their fitness landscape would be more similar to the fitness landscapes we observed at weak and suboptimal promoter activity, where purifying selection against destabilizing mutations would be more intense, thereby preserving protein sequence.

Materials and Methods:

All strains, reagents, compounds and softwares used in this study are listed and referenced in Table S11.

1.- Strains, media and plasmids

MC1061 is the *E. coli* strain used for all cloning and mutagenesis steps, whereas *E. coli* BL21 (DE3) is the one used for the experiments conducted in this study. Transformations with plasmids were done according to standard procedures with in-house chemically competent cells (*38*). Transformed bacteria were grown and selected on 2x YT + glucose medium (1.0 % yeast extract, 1.6 % tryptone, 0.2 % glucose, 0.5 % NaCl and 2 % agar; (*39*)) with ampicillin (AMP; 100 µg/mL). For all experiments

conducted in liquid medium (with the exception of the growth to measure DfrB1 activity; see below), bacteria were grown in LB (Luria-Bertani) medium (0.5 % yeast extract, 1.0 % tryptone, 1.0 % NaCl; (*40*)) with AMP and with or without L(+)-arabinose (0.001 to 0.4 %) and trimethoprim (TMP; 10 µg/mL in DMSO). To measure DfrB1 activity, bacteria were grown in TB (Terrific broth) medium (2.4 % yeast extract, 1.2 % tryptone, 0.4 % glycerol and 89 nM potassium phosphate).

The *dfrB1* gene was expressed in bacteria from a pBAD vector which allows arabinose-controlled induction (*41*). pBAD-*dfrB1* was constructed as follows. The pBAD vector was amplified from pBAD-*chuA* plasmid (*42*) with CLOP198-E9 and CLOP198-F9 (PCR program: 5 min at 95 °C, 35 cycles: 20 s at 98 °C, 15 s at 61 °C, 2 min 30 at 72 °C, and a final extension of 3 min at 72 °C). *dfrB1* was amplified from pDNM plasmid with CLOP198-A10 and CLOP198-B10 (PCR program: 5 min at 95 °C, 5 cycles: 20 s at 98 °C, 15 s at 62 °C, 15 s at 72 °C, 30 cycles: 20 s at 98 °C, 15 s at 72 °C, 15 s at 72 °C, and a final extension of 3 min at 72 °C). Both PCR products were incubated for 1 h at 37 °C with 20 U of DpnI enzyme to remove parental DNA and subsequently purified on magnetic beads (Axygen AxyPrep Mag PCR Clean-up Kit). The vector (pBAD) and insert (*dfrB1*), with their overlapping regions at each extremity, were then assembled by Gibson DNA assembly (*43*), producing pBAD-*dfrB1* plasmid.

To measure by cytometry the effect of mutating the second residue of DfrB1 on the expression level, we fused GFP in 3' of the *dfr*B1 gene sequence and of its mutant. The plasmid was amplified from pBAD-*dfrB1* (constructed above) or from pBAD-*dfrB1*(E2R) (isolated from the DMS plasmid collection - see below) with CLOP273-A3/B3 (PCR program: 5 min at 95 °C, 35 cycles: 20 s at 98 °C, 15 s at 61 °C, 2 min 30 at 72 °C, and a final extension of 3 min at 72 °C). GFP was amplified from pBAD-*sfGFP* (*44*) with CLOP273-D3/F3 (PCR program: 5 min at 95 °C, 5 cycles: 20 s at 98 °C, 15 s at 62 °C, 20 s at 72 °C, 30 cycles: 20 s at 98 °C, 15 s at 72 °C, 20 s at 72 °C, and a final extension of 3 min

at 72 °C). The PCR products corresponding to the plasmids and insert were incubated for 1 h 30 at 37 °C with 10 U of DpnI enzyme to remove parental DNA. Subsequently, the plasmids (pBAD-*dfrB1* and pBAD-*dfrB1*(E2R)) and insert (*sfGFP*) were purified on magnetic beads (Axygen AxyPrep Mag PCR Clean-up Kit) and assembled by Gibson DNA assembly (*43*), producing pBAD-*dfrB1*-*sfGFP* and pBAD-*dfrB1*(E2R)-*sfGFP*.

We also fused the GFP to a fragment corresponding only to DfrB1 first 25 residues. Vectors and the *sfGFP* insert were amplified from the same plasmids as above but using respectively CLOP273-A3/C3 or CLOP273-E3/F3 primers. After DpnI digestion and magnetic beads purification, Gibson DNA assembly produced pBAD-*dfrB1*[1-25]-*sfGFP* and pBAD-*dfrB1*[1-25](E2R)-*sfGFP* plasmids.

To insert specific mutations in the *dfrB1* sequence, we performed site-directed mutagenesis based on the QuickChange™ Site-Directed Mutagenesis System (Stratagene, La Jolla, CA). Briefly, we amplified the pBAD-*dfrB1* plasmid using pairs of primers containing the desired mutation at the center (see Table S12 for the primers specific to each mutation) (PCR program: 2 min at 95 °C, 22 cycles: 20 s at 98 °C, 15 s at 68 °C, 3 min at 72 °C, and a final extension of 5 min at 72 °C). The PCR products were then incubated for 1 h 30 at 37 °C with 6 U of DpnI enzyme to remove parental DNA and mutated plasmids were retrieved directly by transformation in bacteria. We used this method to generate the following mutants: pBAD-*dfrB1*(F24V), pBAD-*dfrB1*(K33E), pBAD-*dfrB1*(K33L), pBAD-*dfrB1*(K33M), pBAD-*dfrB1*(S59Y) and pBAD-*dfrB1*(I68M).

To generate a mutant in which the first methionine codon of DfrB1 is replaced by a stop codon (M1* mutant - pBAD-*dfrB1*(M1*)), we performed a PCR reaction using 10 ng of pBAD-*dfrB1* plasmid as a template (PCR program: 5 min at 95 °C, 25 cycles: 20 s at 98 °C, 15 s at 59 °C, 2 min 30 at 72 °C, and a final extension of 3 min at 72 °C) with the non-overlapping primers (CLOP265-A1 and CLOP259-H5)

designed to incorporate a TGA stop codon on the forward primer to replace the AUG. The PCR product was then digested at 37 °C for 1 hour with 10 U DpnI. The digested PCR product was purified on magnetic beads, quantified and ~50 ng were taken for phosphorylation using 5 U of T4 PNK in T4 ligase buffer for 30 min at 37 °C (final reaction volume of 5 µL). Subsequently, ligation was done in a reaction volume of 10 µL by adjusting the volume of T4 ligase buffer, adding 10 U of T4 DNA ligase and incubating for 1 hour at 22 °C. Finally, mutated plasmid was retrieved directly by transformation in *E. coli* BL21 bacteria.

For fitness validation, some of the plasmids that were used were directly isolated from the mutagenesis pools but confirmed by sequencing. Briefly, an aliquot of glycerol stock for the position of interest was diluted and plated on solid medium. Colony PCR was performed on 32 colonies to amplify the *dfrB1* gene using CLOP228-C1/D1 (PCR program: 5 min at 94 °C, 30 cycles: 30 s at 94 °C, 30 s at 54 °C, 35 s at 72 °C, and a final extension of 3 min at 72 °C). Two subsequent PCR rounds were used to add specific row-column and plate barcodes (*45*). These PCRs are described in detail in the library sequencing section below. All pooled samples were sent to the Genomic Analysis Platform (IBIS, Quebec, Canada) for paired-end 300 bp sequencing on a MiSeq (Illumina). After identification of individual mutations, plasmids for clones of interest were retrieved and purified from bacteria using the Presto Mini Plasmid Kit.

All PCR reactions mentioned above or described below were performed with oligonucleotides defined in (Table S12), using Kapa polymerase at the exception of the colony PCRs for which Taq DNA polymerase was used. The integrity of all assembled and mutagenized plasmids was confirmed by Sanger sequencing (Plateforme de séquençage et de génotypage des génomes, CRCHUL, Canada) using either CLOP194-G11 or CLOP196-B5 as a sequencing primer.

2.- Deep mutational scanning

A single-site mutation library of *dfrB1* was generated by a PCR-based saturation mutagenesis method plasmid with the oligonucleotides defined in (Table S12). Forward oligonucleotides which have degenerate nucleotides (NNN) so each codon of *dfrB1* is mutated were used in combination with a single reverse primer located in the plasmid outside the coding sequence. A first PCR to generate an amplicon containing the desired mutations was conducted following these steps: 5 min at 95 °C, 35 cycles: 20 s at 98 °C, 15 s at 60 °C, 30 s at 72 °C, and a final extension of 1 min at 72 °C. The resulting PCR product was then used as a mega primer to introduce the mutations in pBAD-*dfrB1* by amplifying the whole plasmid (PCR program: 5 min at 95 °C, 22 cycles: 20 s at 98 °C, 15 s at 68 °C, 5 min at 72 °C, and a final extension of 7 min at 72 °C). The long PCR product was digested for 90 min at 37 °C with 6 U of DpnI to remove parental DNA. The digestion products for individual positions were transformed into *E. coli* MC1061. More than a thousand colonies were retrieved from each transformation. After addition of glycerol, libraries for each position were stored separately. From the same pools of bacteria, plasmids were also extracted and purified. PCR amplification and Illumina MiSeq sequencing, performed on these plasmids allowed library quality control assessment (Fig. S2). In an initial round, mutagenesis of position 39 was not successful. Mutagenesis at this position was repeated separately and added to the final library after quality control.

3.- Bulk competition assay

First, *E. coli* BL21 was transformed with 75 ng of each individual position mutant pool (DNA miniprep from the DMS step above). Start and stop codon positions (positions 1 and 79) were omitted. All colonies were retrieved from each transformation plate in 5 mL of 2x YT liquid medium. OD_{600} was measured for each pool (final OD_{600} between 40-80 after being scraped and resuspended in medium). At this step, 15 % glycerol was added to the medium and an aliquot of each pool was stored individually for

further use. In parallel, pools were equally mixed at an OD_{600} of 25 to generate a starting pool for the bulk competition assay. This master pool was stored for further use.

Two separate bulk competition assays were performed. Details concerning arabinose concentrations, the presence or absence of TMP and the number of replicates are indicated in Table S1. Briefly, the master pool was used to inoculate at OD_{600} = 0.01 a first preculture in LB + AMP medium. After an overnight incubation at 37 °C with agitation (250 rpm), cells were diluted 1:100 in fresh medium with the addition of different amounts of arabinose. Following an 18-hour incubation at 37 °C (250 rpm), cells were diluted again at OD_{600} = 0.025 in a final volume of 4 mL of fresh medium containing arabinose with the addition of TMP or DMSO (TMP solvent - no TMP control). Cultures were then incubated as above until OD_{600} reached 0.8 (5 generations). At this point, 125 µL was used to dilute back to OD_{600} = 0.025 in fresh medium and cells were grown until once again, OD_{600} reached 0.8 (another 5 generations). Finally, at two timepoints (18-hour preculture = Timepoint 0, after 10 generations = Timepoint 10), 3 mL of the culture was used to extract plasmid DNA (Fig. S3). These samples were treated as described below to prepare libraries for sequencing.

4.- Single mutant library sequencing

Libraries for sequencing were generated as described in Dionne et al. (*46*). Briefly, three PCR steps were done. The first one was performed directly on the small plasmid DNA preparations (4.5 ng of plasmid) corresponding to single-codon mutant libraries or to bulk competition assay samples (PCR program: 3 min at 98 °C, 20 cycles: 30 s at 98 °C, 15 s at 60 °C, 30 s at 72 °C, and a final extension of 1 min at 72 °C). The second PCR was performed to add Row and Column barcodes (*45*) for identification in a 96-well plate (PCR program: 3 min at 98 °C, 15 cycles: 30 s at 98 °C, 15 s at 60 °C, 30 s at 72 °C, and a final extension of 1 min at 72 °C). For this second PCR, the first PCR product was used as a template (2.25 µL of a 1/2500 dilution). Quantification on gel of the second PCR product using Image Lab

(BioRad Laboratories) allowed us to mix the libraries so each has a roughly equal amount in the final library. Mixed PCRs were purified on magnetic beads and quantified using a NanoDrop (ThermoFisher). Finally, 0.0045 ng of the purified mixed PCRs were used as template for the third PCR which adds a plate barcode as well as Illumina adapters (PCR program: 3 min at 98 °C, 18 cycles: 30 s at 98 °C, 15 s at 61 °C, 35 s at 72 °C, and a final extension of 1 min at 72 °C). Each reaction for the third PCR was performed in quadruplicate and then combined. After purification on magnetic beads, libraries were quantified using a NanoDrop (ThermoFisher) and sent to the Genomic Analysis Platform (IBIS, Quebec, Canada) for paired-end 300 bp sequencing on a MiSeq (Illumina) or to the Plateforme de séquençage et de génotypage des génomes (CRCHUL, Quebec, Canada) for paired-end 250 bp sequencing on a NovaSeq (Illumina). All raw data are available at SRA BioProject PRJNA842350 (accession numbers SRR19419448 and SRR19419449).

5.- Bacterial growth curves

To measure fitness for individual WT or mutant DfrB1, growth was followed with serial OD_{600} measurements in a plate reader. From an overnight preculture grown in LB + AMP medium, cells were diluted 1:100 in fresh medium with the addition of 0 % to 0.4 % arabinose according to the experiment. Following an 18-hour incubation at 37 °C (250 rpm), cells were diluted again at $OD_{600} = 0.01$ in a final volume of 200 µL of fresh medium containing arabinose with the addition of TMP or DMSO. The 96-well plate was incubated at 37 °C in an Infinite M Nano plate reader (Tecan) for 20 hours. OD_{600} measurements were taken every 15 min. Plate was agitated at 200 rpm in between measurements. Growth curves were analyzed with the Growthcurver R package (*47*) to calculate the area under the curve as an estimate of fitness. The percentage of recovered growth upon arabinose induction was calculated as follows:

$$g = 100 * \left(\frac{AUC_{TMP,\, induction}}{AUC_{no\, TMP,\, induction}} \right)$$

where g refers to the recovered growth percentage and *AUC* refers to the area under the curve in the conditions specified by the subscript (with or without TMP).

6.- Expression level measurement by flow cytometry

To measure GFP level by cytometry, a first preculture in LB + AMP medium was grown with bacteria containing the plasmid of interest (pBAD-*sfGFP*, pBAD-*dfrB1*-*sfGFP*, pBAD-*dfrB1*(E2R)-*sfGFP*, pBAD-*dfrB1*[1-25]-*sfGFP* and pBAD-*dfrB1*[1-25](E2R)-*sfGFP*). As for the bulk competition assay, after an overnight incubation at 37 °C with agitation (250 rpm), cells were diluted 1:100 in fresh medium with the addition of different amounts of arabinose. Following an 18-hour incubation at 37 °C (250 rpm), cells were diluted again at OD_{600} = 0.025 in a final volume of 4 mL of fresh medium containing arabinose. Cultures were then incubated as above until OD_{600} reached 0.8 (5 generations). At the different timepoints (18-hour preculture = Timepoint 0, after 5 generations = Timepoint 5), small aliquots of cells were taken and diluted in sterile filtered water to an OD_{600} = 0.05 in 200 µL. GFP fluorescent measurements as well as forward scatter (FSC) and side scatter (SSC) data were collected from a Guava easyCyte HT cytometer (Luminex). From the cytometry data, *E. coli* cells were selected based on FSC and SSC. From the selected data points, GFP fluorescence signal was measured after excitation with a blue laser (wavelength = 488 nm) and detection with a green channel (525/30 nm).

7.- DfrB1 activity measurement

Enzymatic activity of DfrB1 and of different mutants was tested *in vitro*. From a 16-18 h 5 mL LB + AMP preculture (incubation at 37 °C with agitation (230 rpm), a 10 mL culture in TB + AMP was inoculated at OD_{600} = 0.1 and incubated for 3 h at 37 °C (OD_{600} = 0.7 to 1) with agitation. Induction of

expression was initiated by addition of 1 % arabinose and incubation was continued at 22 °C for 16-18 h with agitation. After induction, cells were pelleted for 30 min at 3000 rpm (Eppendorf Centrifuge 5810 R) at 21 °C, resuspended in 300 µL of lysis buffer (0.1 M KH_2PO_4-K_2HPO_4 pH 8.0, 10 mM $MgSO_4$, 1 mM DTT, 0.5 mg/mL lysozyme, 5 U DNase, 1.5 mM benzamidine, 0.25 mM PMSF) and incubated for 2 h at 30 °C with vigorous shaking. Lysates were cleared by a 30 min centrifugation at 3000 rpm (Eppendorf Centrifuge 5810 R) at 21 °C. In a 96-well plate, cleared lysates (20 µL) were combined with 50 µg/mL TMP, 100 µM DHF (synthesized as in (*48*)) and 100 µM NADPH in 50 mM KH_2PO_4-K_2HPO_4 pH 7.0 in a total volume of 100 µL. Measurement of absorbance at 340 nm was taken for 5 min in a Beckman Coulter DTX880 Multimode detector plate reader. Enzyme activity was calculated using the slope for the initial 20 % of substrate consumption.

8.- Data analyses

8.1.- Inferring fitness scores from sequencing data

Quality control of the MiSeq and NovaSeq sequencing data was performed using FastQC version 0.11.4 (*49*). Trimmomatic version 0.39 (*50*) was used to select reads with a minimal length from the raw data (MINLEN parameter: 299 for MiSeq reads, 250 for NovaSeq reads) and trim them to a final length (CROP parameter: 270 for MiSeq reads, 225 for NovaSeq reads) . Selected reads were aligned using Bowtie version 1.3.0 (*51*) to the plate, row, and column barcodes to demultiplex sequences from each pool (arabinose concentration + timepoint). The remaining paired reads were merged with Pandaseq version 2.11 (*52*) and identical sequences were aggregated using vsearch version 2.15.1 (*53*). Finally, aggregated reads were aligned to the reference sequence of DfrB1 to identify the mutant codon in each read. Stop codon UAG was removed from all analysis because it has been shown to have a lower termination efficiency than the other stop codons in *E. coli (54, 55)*.

Raw read counts were normalized by the total number of reads in each pool (arabinose + timepoint). The resulting read proportions were used to calculate selection coefficients based on the end point and the starting point of each experiment using the following equation:

$$s = \frac{log_2\left(\frac{N_{mut,t10} / median(N_{WT,t10})}{N_{mut,t0} / median(N_{WT,t0})}\right)}{k}$$

where s is the selection coefficient, N is the number of reads for the corresponding mutant at a specific timepoint, and k is the number of generations (k = 10). Finally, biases in the bimodal distributions for some samples were corrected using the mixtools R package (*56*) so that the second mode of the distributions would be centered at zero, the corresponding value for neutral mutations.

Since the selection coefficients were correlated with changes in growth recovery (Fig. S7), we used the selection coefficients (Table S3) and the growth recovery for WT at each promoter activity level (Fig. 1C) to derive a pseudo growth recovery metric that could approximate the expected growth recovery for each mutation. This metric was calculated as follows:

$$\hat{g} = g_{WT} * (1 + s)$$

where $\hat{g}$ is the estimated pseudo growth recovery at a given promoter activity level, g_{WT} is the experimentally determined growth recovery for the WT, and *s* is the selection coefficient of a given mutant.

8.2.- Evolutionary analysis (GEMME, Evol, jackhmmer)

A set of seven DfrB1 homologs was extracted from data published by Toulouse et al. (*57*). These seven sequences were then submitted to the online jackhmmer tool (https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer) (*58*) using the Uniprot database as a target to

look for additional homologous sequences. The resulting set of sequences were filtered with the following criteria, considering that the full-length DfrB1 sequence is 78 residues long:

- E-score < 1e-6
- Alignment length > 50
- Sequence length < 100

Applying these filters resulted in a final set of 82 high-confidence DfrB1 homologs. These sequences were aligned with MAFFT version 7.475 (*59*, *60*) using the iterative refinement method (*61*, *62*). The resulting alignment was then analyzed with the Evol library from the ProDy suite version 2.0 (*63*) to calculate Shannon entropy at each position in the alignment as a metric of evolutionary variation. In parallel, the same MAFFT alignment was analyzed with the GEMME model (*19*) to obtain predictions of the fitness landscape based on the variation in the alignment.

8.3.- Machine learning analyses

8.3.1.- k-means clustering

We used k-means clustering to identify the typical patterns of expression-dependent changes in fitness effects (Fig. 2C). k-means clustering was run using the kmeans function from the R base stats package (*64*) by setting the seed and using values of k from 1 to 10 (Fig. S9A) and the following parameters: iter.max = 10, nstart = 25. k = 4 was selected as the best compromise between parsimony (based on the diminishing decrease in the sum of squared errors) and interpretability (cluster visualization in Fig. 2C). Centroids reflecting the central selection coefficients of mutants from each cluster are visualized in Fig. S9B and provided in Table S5. To estimate relative enrichment of mutants from specific clusters at the sites of interest, we organized the data in separate contingency tables indicating whether a particular mutant was assigned or not to that site and its corresponding cluster. We calculated the log2 fold ratio of observed vs expected counts of mutants from each cluster at each site (Fig. S9C). To compare if the

distribution was statistically significant from expectations, we derived contingency tables with counts of mutations from each cluster belonging to the region analyzed and to the rest of the protein. We performed chi-squared tests on these contingency tables and used the Benjamini-Hochberg correction (FDR < 0.5) for multiple hypothesis testing.

8.3.2.- Random forest regressor

A random forest regressor was trained to identify the features that better explain variation in promoter activity level dependent Δs. Explanatory variables included in the model were:

- Relative solvent accessibility (RSA): Solvent accessibility was calculated using DSSP version 2.2.1 (*65*) on the biological assembly of DfrB1 (PDB: 2RK1) (*15*). RSA was then obtained by dividing the solvent accessibility of each residue by the maximum solvent accessibility of that residue as described by Miller, et al. (*66*).
- Biophysical effects of mutations: FoldX version 5.0 (*31*, *67*) was used with the MutateX wrapper (*68*) to simulate all possible mutations on the biological assembly of DfrB1 (PDB: 2RK1). Effects of mutations on subunit stability and binding affinity at the dimerization and tetramerization interfaces were measured.
- Differences in amino acid scores: 57 amino acid indexes were downloaded from the ProtScale database (*26*) on January 21st, 2019, as well as the index on propensity of each amino acid to participate in protein-protein interaction interfaces (*27*). For each mutation, we calculated the differences in index scores between the mutant residue and the WT residue.

All the explanatory variables were divided by their maximum values to scale them between -1 and 1. The random forest model was trained using the sklearn Python library (*69*, *70*) with 80 % of the data and tested with the remaining 20 %. We used the Python rfpimp library (https://github.com/parrt/random-forest-importances) to estimate relative importances of each variable

by calculating the decrease in R^2 of the predictions if the values of a particular feature are permuted (Fig. S16A). Due to the high degree of collinearity in our set of explanatory variables, retraining the random forest with different seeds can suggest a different set of top variables. We introduced a random variable *N(0, 1)* as an internal control in the training set to identify which variables have significant contributions to the model. We selected the variables that have a higher relative importance than the random variable and retrained the random forest to obtain the final model. Finally, we evaluated the relative importance of the variables in the final model by comparing the decrease in R^2 when the final model is retrained without that variable (Fig. S16D).

8.4.- Statistical analyses

8.4.1.- ANOVAs

We used two different types of ANOVAs to characterize expression-dependent changes in fitness effects of coding mutations. First, we performed an ANOVA for each mutation considering all of its replicates at all the promoter activity levels to identify mutants whose fitness effects were significantly affected by changes in gene expression. To correct for multiple hypothesis testing, we applied the Benjamini Hochberg correction with a false discovery rate of 0.05 using the FDRestimation R package (Table S4) (*71*). Then, we did an ANOVA on ranks with all replicates of all mutations at all the promoter activity levels by applying the Aligned Rank Transform from the ARTool R package (*72*, *73*) to estimate the contribution of the interaction between promoter activity level and coding mutants to fitness relative to their separate contributions (Table S6).

We used ANOVAs to analyze differences in the distribution of Δs based on the protein sites (Fig. 5C) and differences in the fluorescence measured for different constructs (Fig. 4C). In both cases, we complemented the ANOVA with Tukey's honest significant difference post-hoc test using the agricolae R package (*74*).

References

1. J. A. G. M. de Visser, J. Krug, Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490 (2014).

2. I. Fragata, A. Blanckaert, M. A. Dias Louro, D. A. Liberles, C. Bank, Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82 (2019).

3. B. Lehner, Molecular mechanisms of epistasis within and between genes. *Trends Genet.* **27**, 323–331 (2011).

4. M. C. King, A. C. Wilson, Evolution at two levels in humans and chimpanzees. *Science*. **188**, 107–116 (1975).

5. P. J. Wittkopp, G. Kalay, Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69 (2011).

6. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*. **312**, 111–114 (2006).

7. K. M. Brown, M. A. Depristo, D. M. Weinreich, D. L. Hartl, Temporal constraints on the incorporation of regulatory mutants in evolutionary pathways. *Mol. Biol. Evol.* **26**, 2455–2462 (2009).

8. X. Li, J. Lalić, P. Baeza-Centurion, R. Dhar, B. Lehner, Changes in gene expression predictably shift and switch genetic interactions. *Nat. Commun.* **10**, 3886 (2019).

9. L. Jiang, P. Mishra, R. T. Hietpas, K. B. Zeldovich, D. N. A. Bolon, Latent effects of Hsp90 mutants revealed at reduced expression levels. *PLoS Genet.* **9**, e1003600 (2013).

10. S. E. Castel, A. Cervera, P. Mohammadi, F. Aguet, F. Reverter, A. Wolman, R. Guigo, I. Iossifov, A. Vasileva, T. Lappalainen, Modified penetrance of coding variants by cis-regulatory variation contributes to disease risk. *Nat. Genet.* **50**, 1327–1334 (2018).

11. K. H. Pattishall, J. Acar, J. J. Burchall, F. W. Goldstein, R. J. Harvey, Two distinct types of trimethoprim-resistant dihydrofolate reductase specified by R-plasmids of different compatibility groups. *J. Biol. Chem.* **252**, 2319–2323 (1977).

12. E. E. Howell, Searching sequence space: two different approaches to dihydrofolate reductase catalysis. *Chembiochem*. **6**, 590–600 (2005).

13. M. Faltyn, B. Alcock, A. McArthur, Evolution and nomenclature of the trimethoprim resistant dihydrofolate (dfr) reductases. *Preprints* (2019) (available at https://www.preprints.org/manuscript/201905.0137).

14. N. Narayana, D. A. Matthews, E. E. Howell, N.-H. Xuong, A plasmid-encoded dihydrofolate reductase from trimethoprim-resistant bacteria has a novel D2-symmetric active site. *Nat. Struct. Mol. Biol.* **2**, 1018–1025 (1995).

15. J. M. Krahn, M. R. Jackson, E. F. DeRose, E. E. Howell, R. E. London, Crystal structure of a type

II dihydrofolate reductase catalytic ternary complex. *Biochemistry*. **46**, 14878–14888 (2007).

16. C. Lemay-St-Denis, S.-S. Diwan, J. N. Pelletier, The Bacterial Genomic Context of Highly Trimethoprim-Resistant DfrB Dihydrofolate Reductases Highlights an Emerging Threat to Public Health. *Antibiotics (Basel)*. **10** (2021), doi:10.3390/antibiotics10040433.

17. H. Kacser, J. A. Burns, The molecular basis of dominance. *Genetics*. **97**, 639–666 (1981).

18. S. Stefl, H. Nishi, M. Petukh, A. R. Panchenko, E. Alexov, Molecular mechanisms of disease-causing missense mutations. *J. Mol. Biol.* **425**, 3919–3936 (2013).

19. E. Laine, Y. Karami, A. Carbone, GEMME: A Simple and Fast Global Epistatic Model Predicting Mutational Effects. *Mol. Biol. Evol.* **36**, 2604–2619 (2019).

20. J. Feng, J. Grubbs, A. Dave, S. Goswami, C. G. Horner, E. E. Howell, Radical redesign of a tandem array of four R67 dihydrofolate reductase genes yields a functional, folded protein possessing 45 substitutions. *Biochemistry*. **49**, 7384–7392 (2010).

21. L. J. Reece, R. Nichols, R. C. Ogden, E. E. Howell, Construction of a synthetic gene for an R-plasmid-encoded dihydrofolate reductase and studies on the role of the N-terminus in the protein. *Biochemistry*. **30**, 10895–10904 (1991).

22. A. R. Schmitzer, F. Lépine, J. N. Pelletier, Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: creating alternative functional configurations. *Protein Eng. Des. Sel.* **17**, 809–819 (2004).

23. M. Verma, J. Choi, K. A. Cottrell, Z. Lavagnino, E. N. Thomas, S. Pavlovic-Djuranovic, P. Szczesny, D. W. Piston, H. S. Zaher, J. D. Puglisi, S. Djuranovic, A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **10**, 5774 (2019).

24. M. B. Strader, R. Derike Smiley, L. G. Stinnett, N. C. VerBerkmoes, E. E. Howell, Role of S65, Q67, I68, and Y69 Residues in Homotetrameric R67 Dihydrofolate Reductase. *Biochemistry*. **40**, 11344–11352 (2001).

25. L. Breiman, Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

26. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch, "Protein Identification and Analysis Tools on the ExPASy Server" in *The Proteomics Protocols Handbook*, J. M. Walker, Ed. (Humana Press, Totowa, NJ, 2005), pp. 571–607.

27. E. D. Levy, S. De, S. A. Teichmann, Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20461–20466 (2012).

28. Z. Wu, X. Cai, X. Zhang, Y. Liu, G.-B. Tian, J.-R. Yang, X. Chen, Expression level is a major modifier of the fitness landscape of a protein coding gene. *Nature Ecology & Evolution*. **6** (2022), pp. 103–115.

29. P. C. Després, A. F. Cisneros, E. M. M. Alexander, R. Sonigara, C. Gagné-Thivierge, A. K. Dubé, C. R. Landry, Asymmetrical dose responses shape the evolutionary trade-off between antifungal resistance and nutrient use. *Nat Ecol Evol*. **6**, 1501–1515 (2022).

30. F. Duveau, D. C. Yuan, B. P. H. Metzger, A. Hodgins-Davis, P. J. Wittkopp, Effects of mutation and selection on plasticity of a promoter activity in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E11218–E11227 (2017).

31. R. Guerois, J. E. Nielsen, L. Serrano, Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387 (2002).

32. D. R. Usmanova, N. S. Bogatyreva, J. A. Bernad, A. A. Eremina, A. A. Gorshkova, G. M. Kanevskiy, L. R. Lonishin, A. V. Meister, A. G. Yakupova, F. A. Kondrashov, D. N. Ivankov, Self-consistency test reveals systematic bias in programs for prediction change of stability upon mutation. *Bioinformatics*. **34**, 3653–3658 (2018).

33. J. Hausser, A. Mayo, L. Keren, U. Alon, Central dogma rates and the trade-off between precision and economy in gene expression. *Nat. Commun.* **10**, 68 (2019).

34. S. Karve, P. Dasmeh, J. Zheng, A. Wagner, Low protein expression enhances phenotypic evolvability by intensifying selection on folding stability. *Nat Ecol Evol*. **6**, 1155–1164 (2022).

35. D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 14338–14343 (2005).

36. C. Bédard, A. F. Cisneros, D. Jordan, C. R. Landry, Correlation between protein abundance and sequence conservation: what do recent experiments say? *Curr. Opin. Genet. Dev.* **77**, 101984 (2022).

37. J.-F. Gout, D. Kahn, L. Duret, Paramecium Post-Genomics Consortium, The relationship among gene expression, the evolution of gene dosage, and the rate of protein evolution. *PLoS Genet.* **6**, e1000944 (2010).

38. R. Green, E. J. Rogers, Transformation of chemically competent E. coli. *Methods Enzymol.* **529**, 329–336 (2013).

39. Cold Spring Harbor Protocols, 2× YT Medium (2014).

40. Cold Spring Harbor Protocols, LB (Luria-Bertani) liquid medium (2006).

41. L. M. Guzman, D. Belin, M. J. Carson, J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).

42. E. C. Hagan, H. L. T. Mobley, Haem acquisition is facilitated by a novel receptor Hma and required by uropathogenic Escherichia coli for kidney infection. *Mol. Microbiol.* **71**, 79–91 (2009).

43. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. **6**, 343–345 (2009).

44. S. J. Miyake-Stoner, C. A. Refakis, J. T. Hammill, H. Lusic, J. L. Hazen, A. Deiters, R. A. Mehl, Generating Permissive Site-Specific Unnatural Aminoacyl-tRNA Synthetases. *Biochemistry*. **49** (2010), pp. 1667–1677.

45. N. Yachie, E. Petsalaki, J. C. Mellor, J. Weile, Y. Jacob, M. Verby, S. B. Ozturk, S. Li, A. G. Cote,

R. Mosca, J. J. Knapp, M. Ko, A. Yu, M. Gebbia, N. Sahni, S. Yi, T. Tyagi, D. Sheykhkarimli, J. F. Roth, C. Wong, L. Musa, J. Snider, Y.-C. Liu, H. Yu, P. Braun, I. Stagljar, T. Hao, M. A. Calderwood, L. Pelletier, P. Aloy, D. E. Hill, M. Vidal, F. P. Roth, Pooled‐matrix protein interaction screens using Barcode Fusion Genetics. *Mol. Syst. Biol.* **12**, 863 (2016).

46. U. Dionne, É. Bourgault, A. K. Dubé, D. Bradley, F. J. M. Chartier, R. Dandage, S. Dibyachintan, P. C. Després, G. D. Gish, N. T. H. Pham, M. Létourneau, J.-P. Lambert, N. Doucet, N. Bisson, C. R. Landry, Protein context shapes the specificity of SH3 domain-mediated interactions in vivo. *Nat. Commun.* **12**, 1597 (2021).

47. K. Sprouffske, A. Wagner, Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*. **17**, 172 (2016).

48. R. L. Blakley, Crystalline Dihydropteroylglutamic Acid. *Nature*. **188**, 231–232 (1960).

49. S. Andrews, FastQC A Quality Control tool for High Throughput Sequence Data (2010), (available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

50. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**, 2114–2120 (2014).

51. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

52. A. P. Masella, A. K. Bartram, J. M. Truszkowski, D. G. Brown, J. D. Neufeld, PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics*. **13**, 31 (2012).

53. T. Rognes, T. Flouri, B. Nichols, C. Quince, F. Mahé, VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. **4**, e2584 (2016).

54. E. B. Kramer, P. J. Farabaugh, The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. *RNA*. **13**, 87–96 (2007).

55. G. Korkmaz, M. Holm, T. Wiens, S. Sanyal, Comprehensive analysis of stop codon usage in bacteria and its correlation with release factor abundance. *J. Biol. Chem.* **289**, 30334–30342 (2014).

56. T. Benaglia, D. Chauveau, D. R. Hunter, D. S. Young, mixtools: an R package for analyzing mixture models. *J. Stat. Softw.* **32** (2009).

57. J. L. Toulouse, T. J. Edens, L. Alejaldre, A. R. Manges, J. N. Pelletier, Integron-Associated DfrB4, a Previously Uncharacterized Member of the Trimethoprim-Resistant Dihydrofolate Reductase B Family, Is a Clinically Identified Emergent Source of Antibiotic Resistance. *Antimicrob. Agents Chemother.* **61**, e02665–16 (2017).

58. L. S. Johnson, S. R. Eddy, E. Portugaly, Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics*. **11**, 431 (2010).

59. K. Katoh, K. Misawa, K.-I. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

60. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements

in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

61. M. P. Berger, P. J. Munson, A novel randomized iterative strategy for aligning multiple protein sequences. *Comput. Appl. Biosci.* **7**, 479–484 (1991).

62. O. Gotoh, Optimal alignment between groups of sequences and its application to multiple sequence alignment. *Comput. Appl. Biosci.* **9**, 361–370 (1993).

63. A. Bakan, A. Dutta, W. Mao, Y. Liu, C. Chennubhotla, T. R. Lezon, I. Bahar, Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics*. **30**, 2681–2683 (2014).

64. R Core Team, R: A Language and Environment for Statistical Computing (2013), (available at http://www.R-project.org/).

65. W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. **22**, 2577–2637 (1983).

66. S. Miller, J. Janin, A. M. Lesk, C. Chothia, Interior and surface of monomeric proteins. *J. Mol. Biol.* **196**, 641–656 (1987).

67. J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, L. Serrano, The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–W388 (2005).

68. M. Tiberti, T. Terkelsen, K. Degn, L. Beltrame, T. C. Cremers, I. da Piedade, M. Di Marco, E. Maiani, E. Papaleo, MutateX: an automated pipeline for in silico saturation mutagenesis of protein structures and structural ensembles. *Brief. Bioinform.* **23**, bbac074 (2022).

69. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, Others, Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. **12**, 2825–2830 (2011).

70. G. Van Rossum, F. L. Drake, *Python 3 Reference Manual* (CreateSpace Independent Publishing Platform, 2009).

71. M. Murray, J. Blume, FDRestimation: Estimate, Plot, and Summarize False Discovery Rates (2020), (available at https://CRAN.R-project.org/package=FDRestimation).

72. J. O. Wobbrock, L. Findlater, D. Gergle, J. J. Higgins, The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '11)*, 143–146 (2011).

73. M. Kay, L. A. Elkin, J. J. Higgins, J. O. Wobbrock, ARTool: Aligned Rank Transform for Nonparametric Factorial ANOVAs (2021), , doi:10.5281/zenodo.594511.

74. Felipe de Mendiburu, Muhammad Yaseen, agricolae: Statistical Procedures for Agricultural Research (2020).

75. J. Dam, T. Rose, M. E. Goldberg, A. Blondel, Complementation between dimeric mutants as a probe of dimer-dimer interactions in tetrameric dihydrofolate reductase encoded by R67 plasmid of E. coli. *J. Mol. Biol.* **302**, 235–250 (2000).

76. L. L. C. Schrödinger, The PyMOL Molecular Graphics System, Version 1.8 (2015).

Acknowledgements: We thank C. Lemay-St.Denis, D. I. Ascencio, R. Durand, M. Dion, F. Mattenberger and the members of the Landrylab for feedback and help on the project and M. Couture for plasmids. Figures 1A-B and S3 were generated using Biorender.

Funding:

Canadian Institutes of Health Research (CIHR) Foundation grant to 387697 (CRL)

NSERC discovery grant RGPIN-N-2018-04686 (JNP)

Canada Research Chair in Cellular Systems and Synthetic Biology (CRL)

Canada Research Chair in Engineering of Applied Proteins (JNP)

FRQNT Merit Scholarship Program for Foreign Studies (PBEEE) (AFC)

Joint funding from MEES and AMEXCID (AFC)

Vanier graduate scholarship (PCD)

Author contributions:

Conceptualization: AFC, IGA, CRL

Data curation: AFC, IGA, KL

Investigation: IGA, AKD, KL and AFC

Formal analysis: AFC, IGA and CRL

Software: AFC, PCD, and PK

Writing – original draft: AFC, CRL, and IGA

Writing – review & editing: AFC, CRL, IGA, PCD, AKD, KL, JNP

Supervision: JNP and CRL

Funding acquisition: JNP and CRL

Methodology: AFC, IGA, CRL, KL, JNP

Project administration: CRL

Resources: CRL

Validation: AFC, IGA

Visualization: AFC

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: Raw sequencing data are available at SRA BioProject PRJNA842350 (accession numbers SRR19419448 and SRR19419449). Code for analysis is available on Zenodo (https://doi.org/10.5281/zenodo.7250327) and Github at: https://github.com/Landrylab/DfrB1_DMS_2022

All data, code, and materials used in the analysis are available in the supplementary files, and dataset on through SRA. All the material is available upon request or from Addgene for the pBAD-*sfGFP* plasmid.

Supplementary Figures:

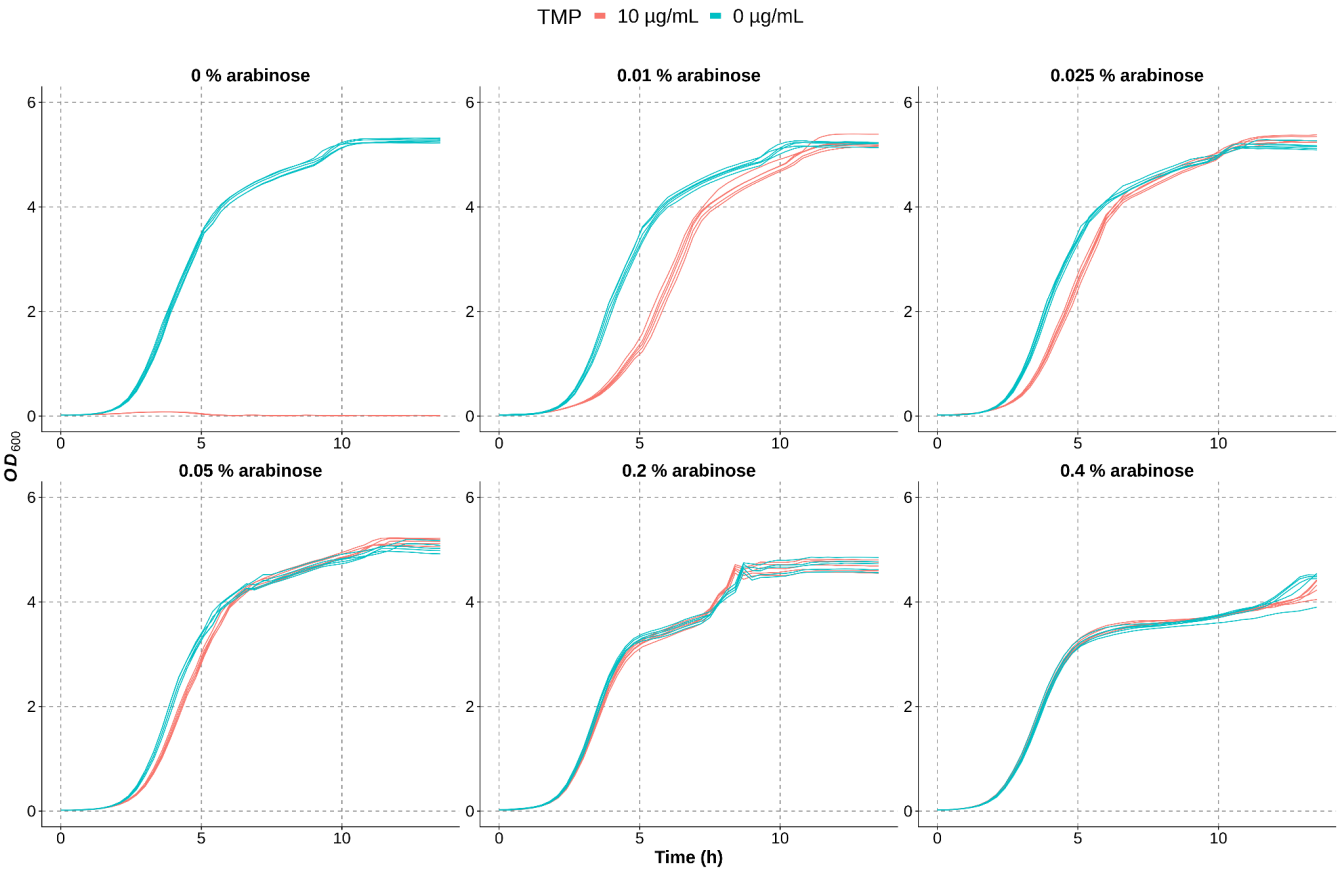


Fig. S1. Dose-dependent growth recovery upon ecDHFR inhibition and DfrB1 expression. Cells expressing WT DfrB1 under the control of the arabinose promoter were grown in LB medium with and without trimethoprim (TMP) and at different concentrations of arabinose as indicated above each panel. Same data as in Fig. 1C. Growth recovery of the inhibition of the endogenous ecDHFR by TMP through the expression of the DfrB1 can be calculated as the difference in the area under the curve for a given arabinose concentration.

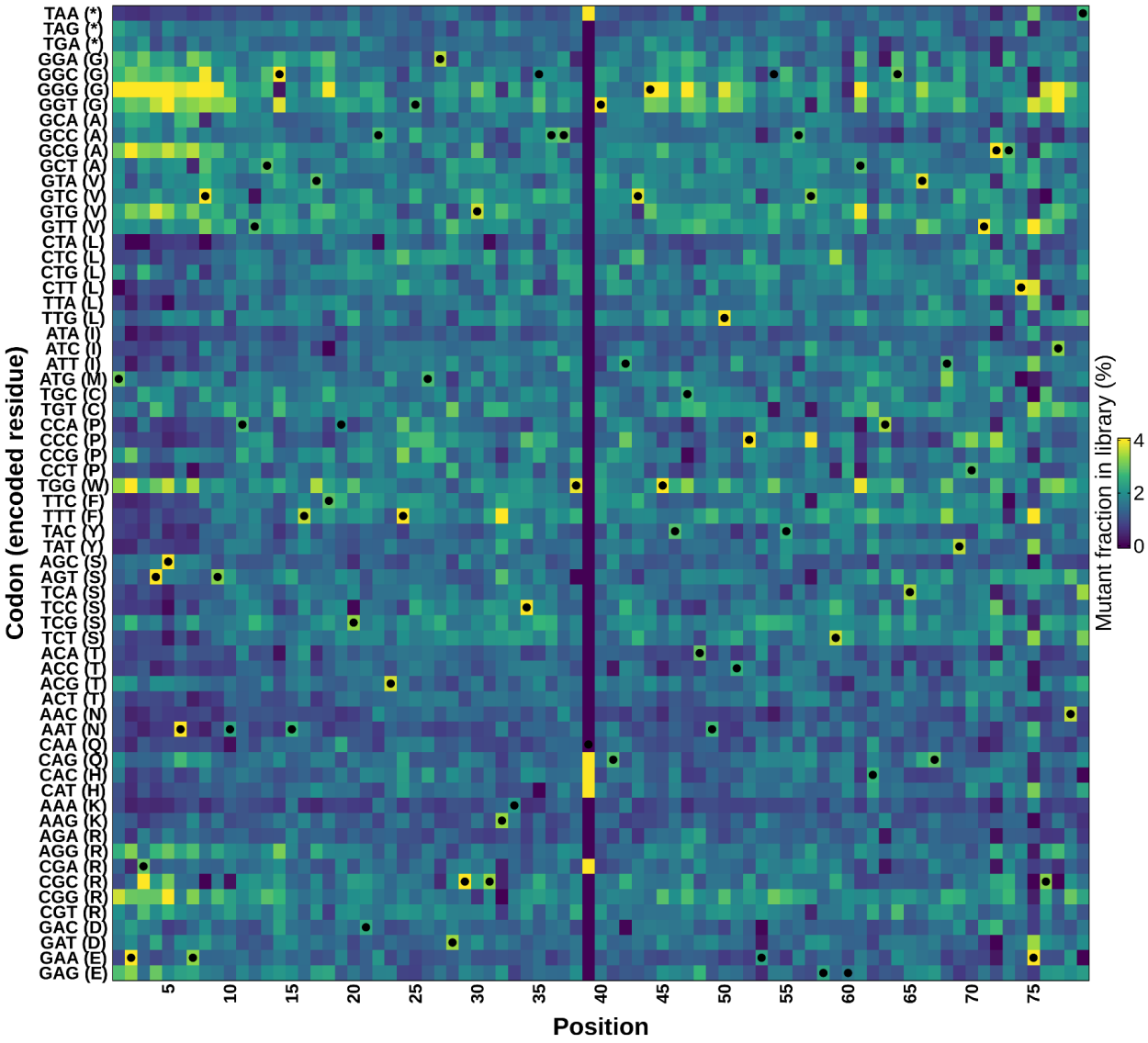


Fig. S2. Quality control of the DfrB1 Deep Mutational Scanning (DMS) library. The libraries corresponding to each position (NNN codons) were pooled and sequenced on a MiSeq PE300 (268 671 reads) to verify the coverage of the various codons across positions. A single position was not covered and was repeated and added to the final library. The color scale represents the percentage of reads that mapped to a mutant out of the library for that position. WT codons are labeled with dots.

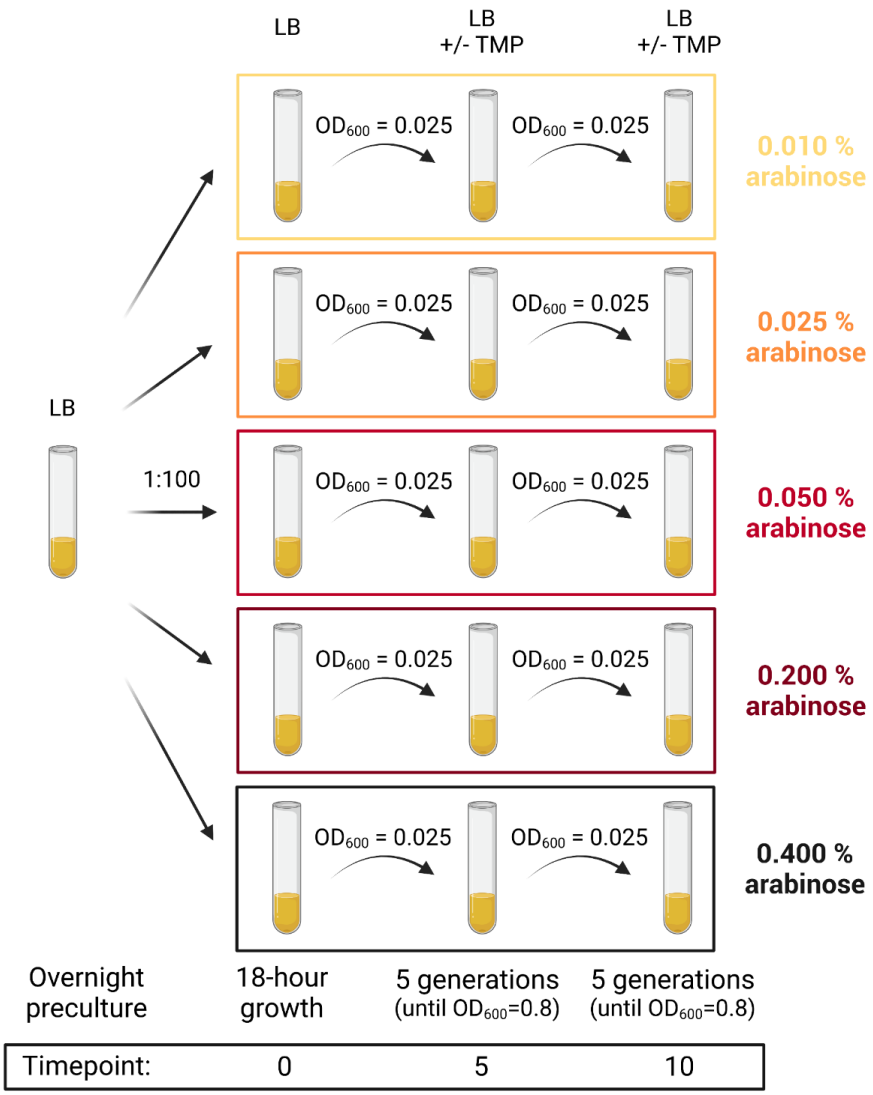
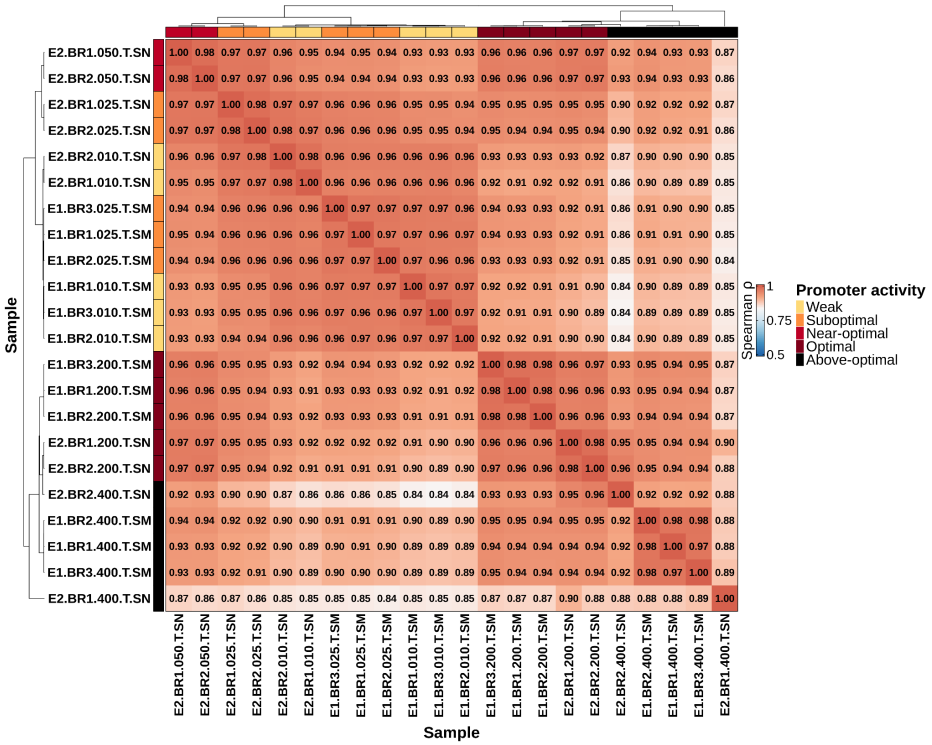


Fig. S3. Experimental setup for the bulk competition assays. Samples for the bulk competition experiments were taken from an overnight preculture of cells transformed with the DMS plasmid library. Cells were grown for five generations in LB medium with different concentrations of arabinose and with or without trimethoprim (TMP) until they reached an optical density of 0.8. They were then diluted again and grown for five more generations. DNA was extracted and sequenced at t = 0 and t = 10 generations. This experiment was repeated twice (see Table S1 for details).



Experiment with TMP	ID	Biological replicates	Arabinose concentration (%)	Sequencing platform
1	E1.BR1.010.T.SM	1	0.01	MiSeq
	E1.BR2.010.T.SM	2		
	E1.BR3.010.T.SM	3		
	E1.BR1.025.T.SM	1	0.025	MiSeq
	E1.BR2.025.T.SM	2		
	E1.BR3.025.T.SM	3		
	E1.BR1.200.T.SM	1	0.2	MiSeq
	E1.BR2.200.T.SM	2		
	E1.BR3.200.T.SM	3		
	E1.BR1.400.T.SM	1	0.4	MiSeq
	E1.BR2.400.T.SM	2		
	E1.BR3.400.T.SM	3		
2	E2.BR1.010.T.SN	1	0.01	NovaSeq
	E2.BR2.010.T.SN	2		
	E2.BR1.025.T.SN	1	0.025	NovaSeq
	E2.BR2.025.T.SN	2		
	E2.BR1.050.T.SN	1	0.05	NovaSeq
	E2.BR2.050.T.SN	2		
	E2.BR1.200.T.SN	1	0.2	NovaSeq
	E2.BR2.200.T.SN	2		
	E2.BR1.400.T.SN	1	0.4	NovaSeq
	E2.BR2.400.T.SN	2		

Fig. S4. Replicates of the bulk competition experiment with selection for DfrB1 activity (with TMP) correlate strongly. Spearman correlation between selection coefficients was estimated for different replicates of the experiment with selection for DfrB1 activity using the antibiotic TMP. Samples were named according to the table (bottom) extracted from Table S1.

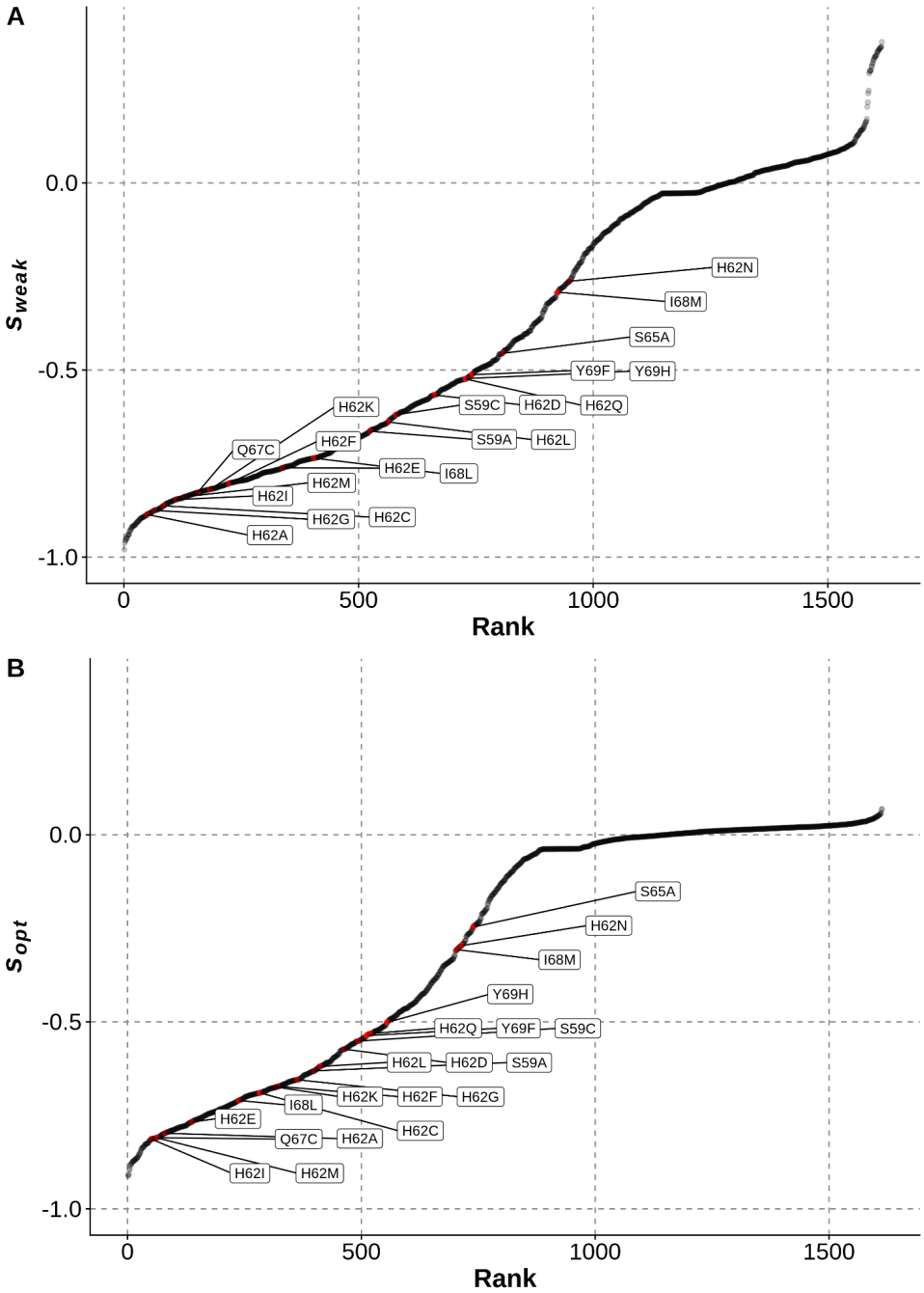


Fig. S5. Selection coefficients recapitulate previously characterized mutants with poor catalytic activity. Ranks and selection coefficients measured by bulk competition assay and sequencing for mutants constructed by Dam et al. (*75*) and Strader et al. (*24*) at a weak promoter activity level for WT DfrB1 (**A**) and at the optimal promoter activity level for WT DfrB1 (B). All mutants had lower activity than the WT in the experimental papers and have negative *s* here, as expected.

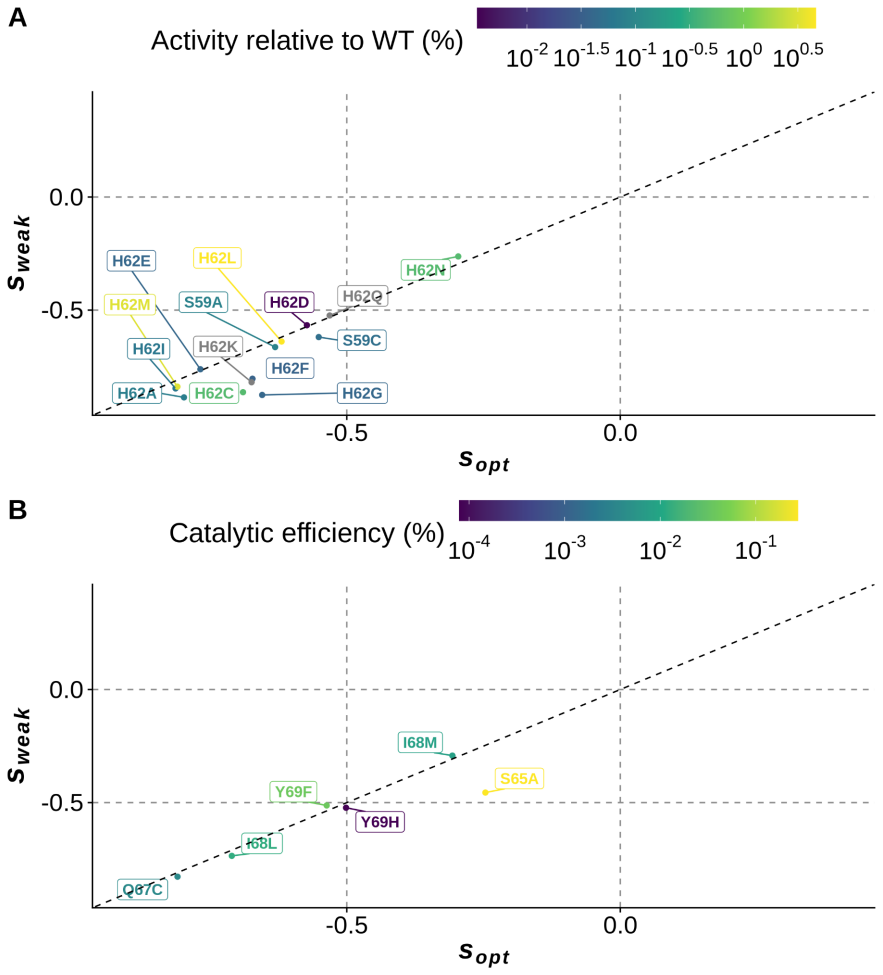


Fig. S6. The fitness effects of only some of the low activity mutants are improved at optimal promoter activity. (**A**) Differences in selection coefficients observed at weak and optimal promoter activity on mutants at the tetramerization interface from Dam et al. (*75*) with respect to their measured activity relative to WT. (B) Differences in selection coefficients observed at weak and optimal promoter activity on active site mutants from Strader et al. (*24*) with respect to their measured catalytic efficiency.

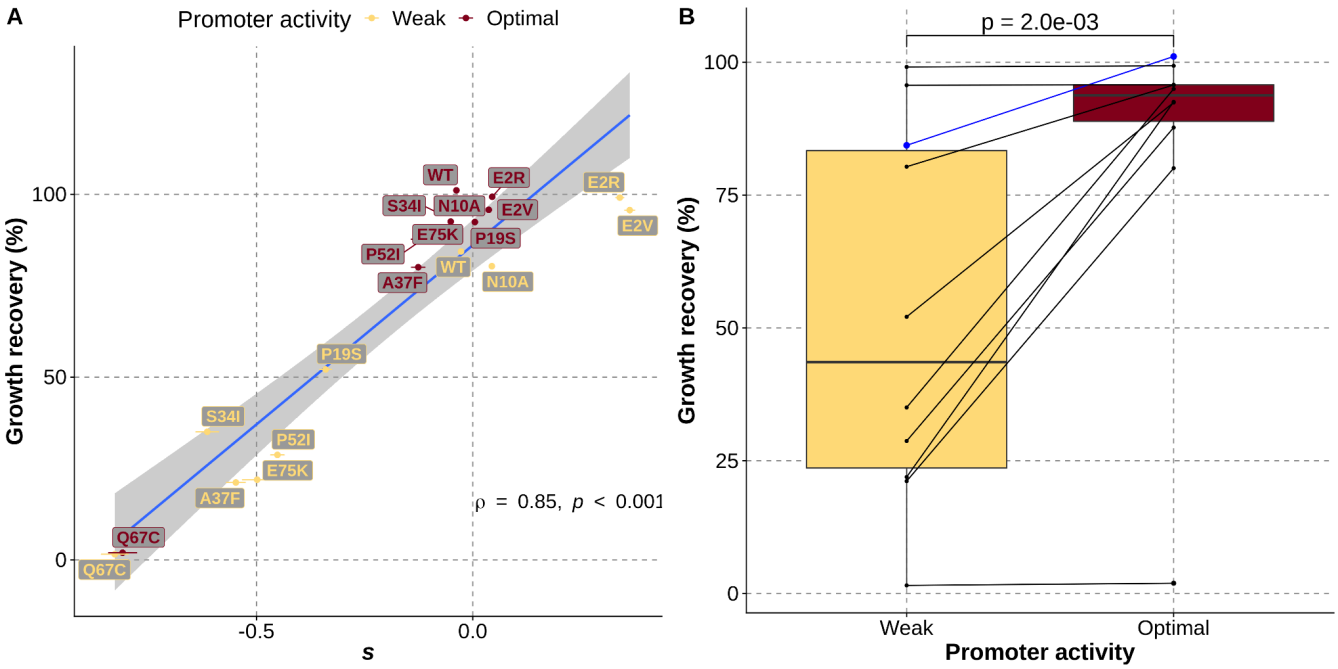


Fig. S7. The effects of mutations measured by bulk competition are recapitulated with individual growth measurements. Mutants with different magnitudes of changes in fitness effects across concentrations of arabinose were selected for validation in individual growth experiments. Growth was measured as the average of the area under the curve for two replicates per mutant at each of two promoter activity levels and subtracted to measure the % of growth recovered upon expression of the DfrB1. (**A**) Growth recovery of individual mutants correlates strongly with selection coefficients. The corresponding mutants and the WT are indicated next to each data point and the colors indicate promoter activities. (B) Promoter activity-level dependent differences in growth rate for the validated mutants on the left. WT is indicated in blue. P-values were calculated using Wilcoxon's test for differences in means of paired samples.

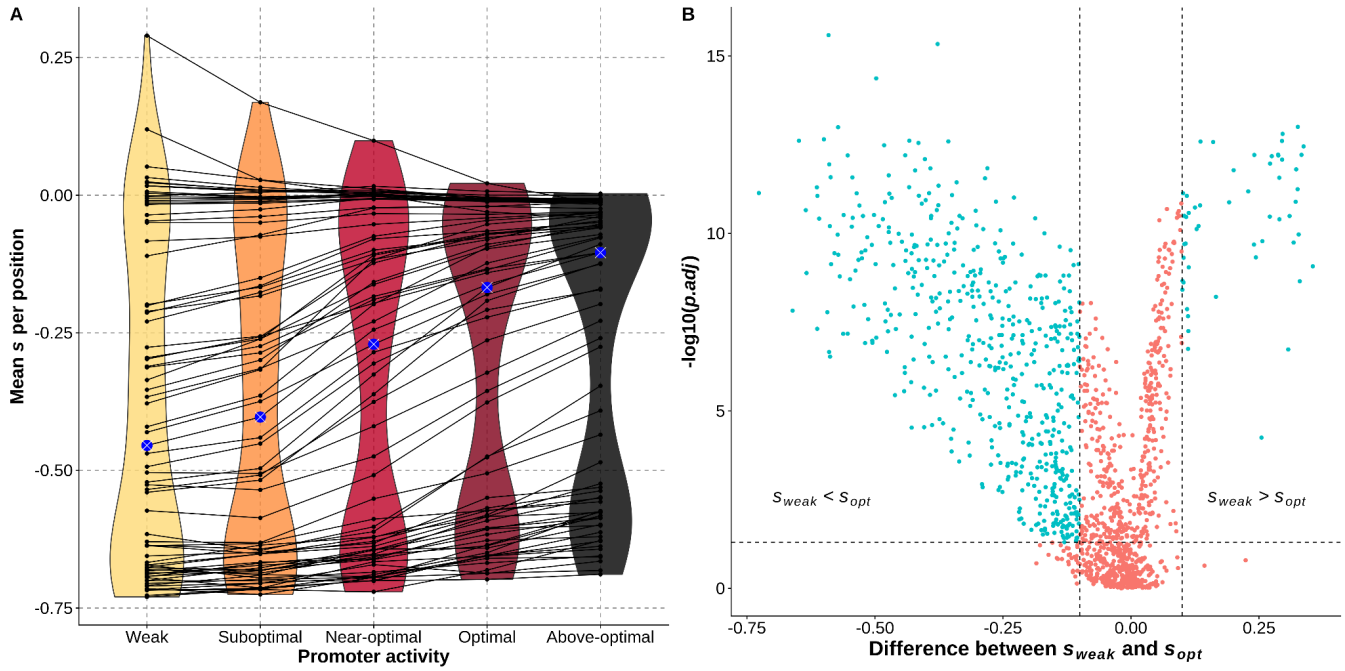


Fig. S8. Significance and magnitude of effects of promoter activity level on the fitness of mutants. (**A**) Averages of selection coefficients observed at each position of the DfrB1 coding sequence. The larger dot indicates the position with the median effect for each promoter activity level. Medians of each distribution are represented with the blue dots with white crosses. (B) For each mutant, we performed an ANOVA testing for significant effects of promoter activity on fitness effects and corrected for multiple hypothesis testing by applying the Benjamini-Hochberg correction with a false discovery rate of 0.05 using the FDR estimation package (*71*). The x-axis shows the difference between s_{weak} and s_{opt} (s_{weak} - s_{opt}) and the y-axis shows the adjusted p-values from the statistical analysis. The horizontal dashed line indicates the cutoff for significance of adjusted p-values ($p < 0.05$) and the vertical lines indicate an arbitrary threshold for magnitude of changes in selection coefficients (greater than 0.1) due to promoter activity level.

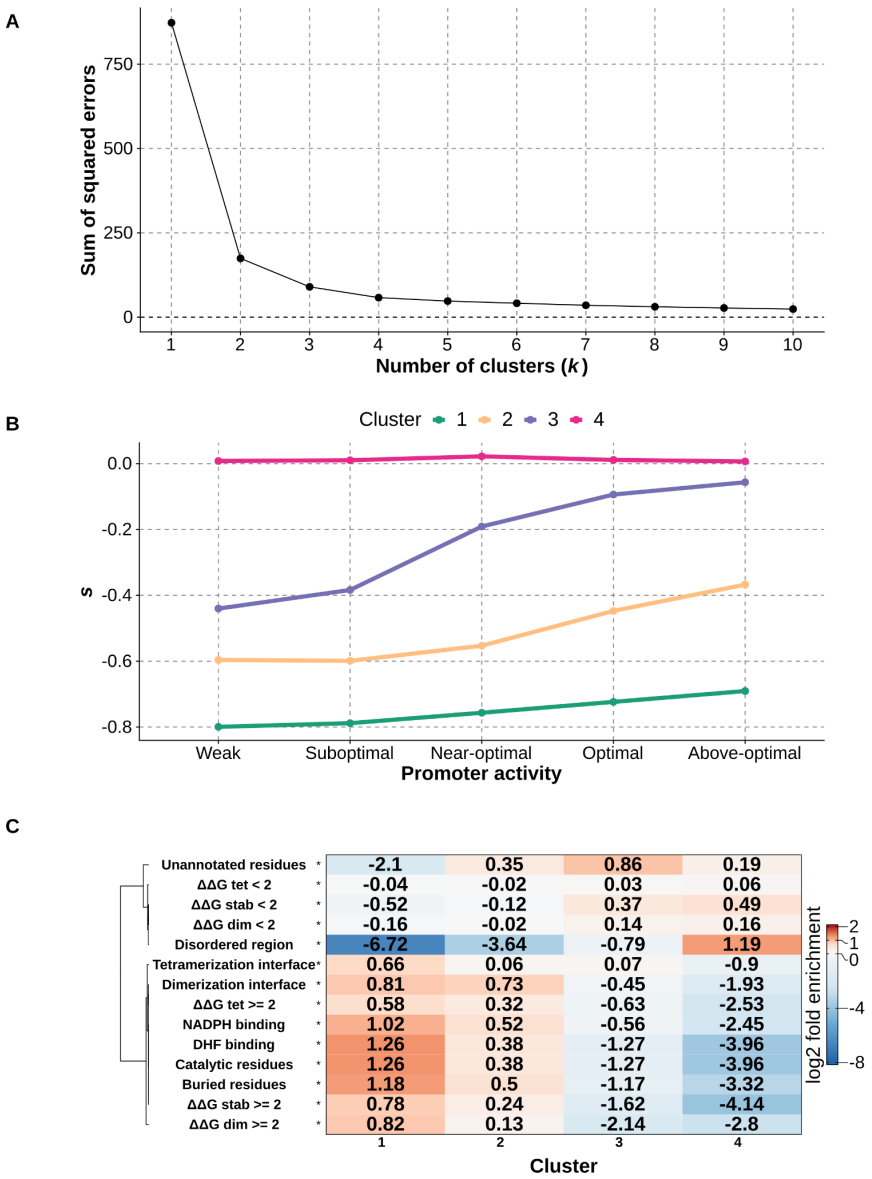


Fig. S9. k-means clusters capture the main patterns of promoter activity-dependent changes in fitness effects. (**A**) Sum of squared errors as a function of the numbers used in k-means clustering. Adding more clusters results in more accurate clustering but increases the probability of overfitting the model. We selected k = 4 because adding more than 4 clusters does not lead to significant decreases in the sum of squared errors. (B) Selection coefficients at each promoter activity level for the centroids of the four clusters. (C) Relative enrichment of each of the four clusters with mutants at particular protein sites and with specific destabilizing effects. Stars indicate significant deviations from expectations as measured with a chi-squared test, adjusted for FDR ($p < 0.05$) with the Benjamini-Hochberg correction.

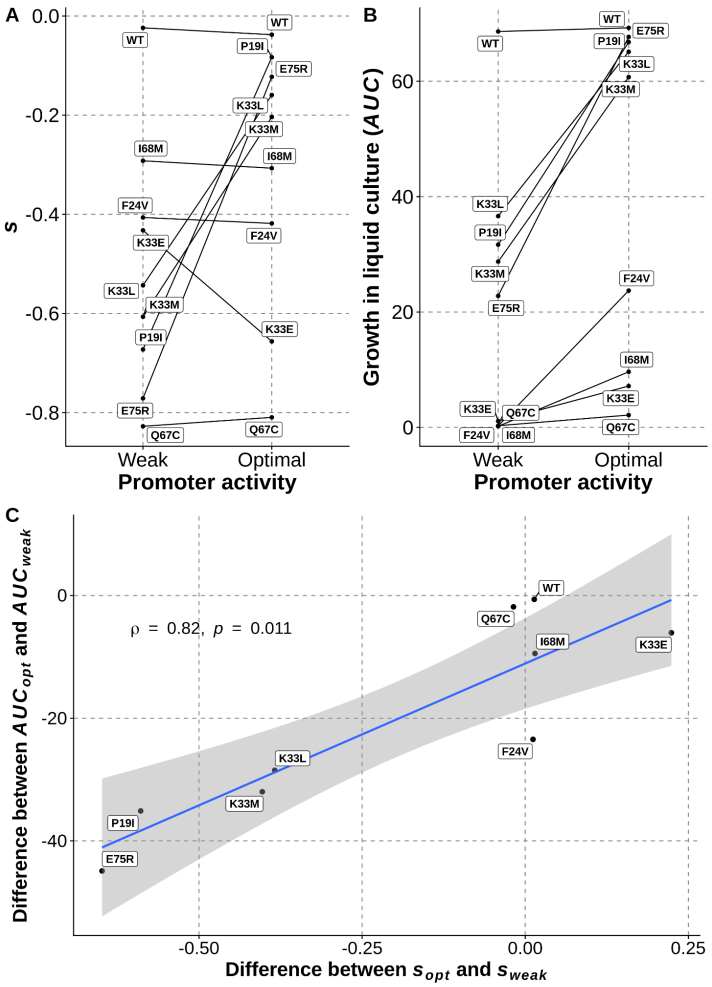


Fig. S10. Rank epistasis is not completely recapitulated in validation cultures. (**A**) Selection coefficients calculated from the bulk competition experiment for a selected group of mutations at weak and optimal promoter activity. Rank epistasis is observed as crossing lines, indicating that a mutation that is more deleterious in a particular condition becomes more beneficial in a different condition. (B) Area under the curve observed for validation cultures of the same selected group of mutations. Some patterns of the rank epistasis observed in the bulk competition experiment are recapitulated but not all of them. (C) The difference in area under the curve in the validation cultures generally correlates with the difference in selection coefficients from the bulk competition experiment. ΔAUC_{weak} was calculated as the differences in the area under the curve observed for growth in liquid cultures at weak and optimal promoter activity levels (AUC_{weak} - AUC_{opt}). Δs_{weak} was calculated as the differences in selection

coefficients observed for each mutant in the bulk competition experiments at weak and optimal promoter activity levels (s_{weak} - s_{opt}).

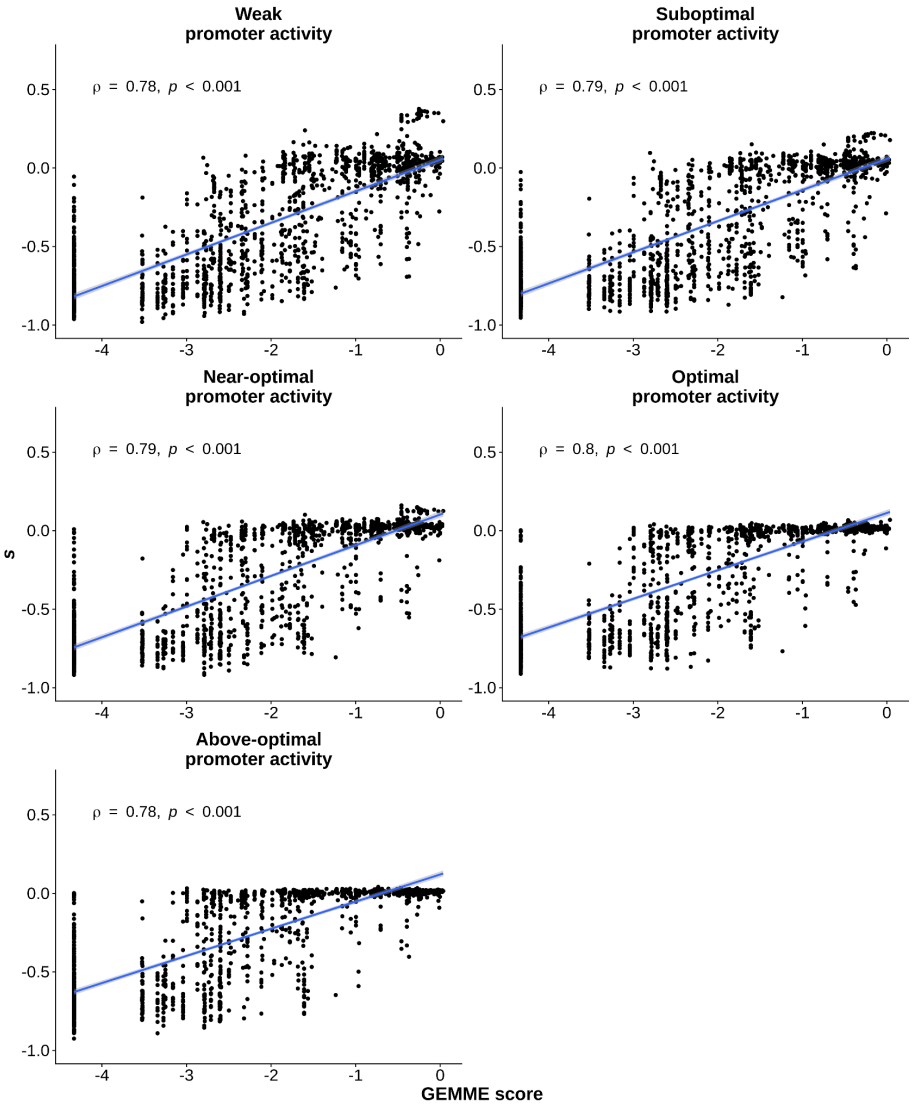


Fig. S11. Selection coefficients measured in the DMS bulk competition experiment with TMP correlate well with fitness effects deduced variation observed in natural sequences. We used GEMME (*19*) to predict mutational effects based on patterns of conservation and substitution. GEMME scores (0 for residues observed as often as the WT DfrB1 residue at that position and negative for residues observed less often) correlate well with the effects measured in the bulk competition assay with the mutant library. These analyses show that the mutational effects measured experimentally reflect the selective pressure these homologous sequences experience in nature.

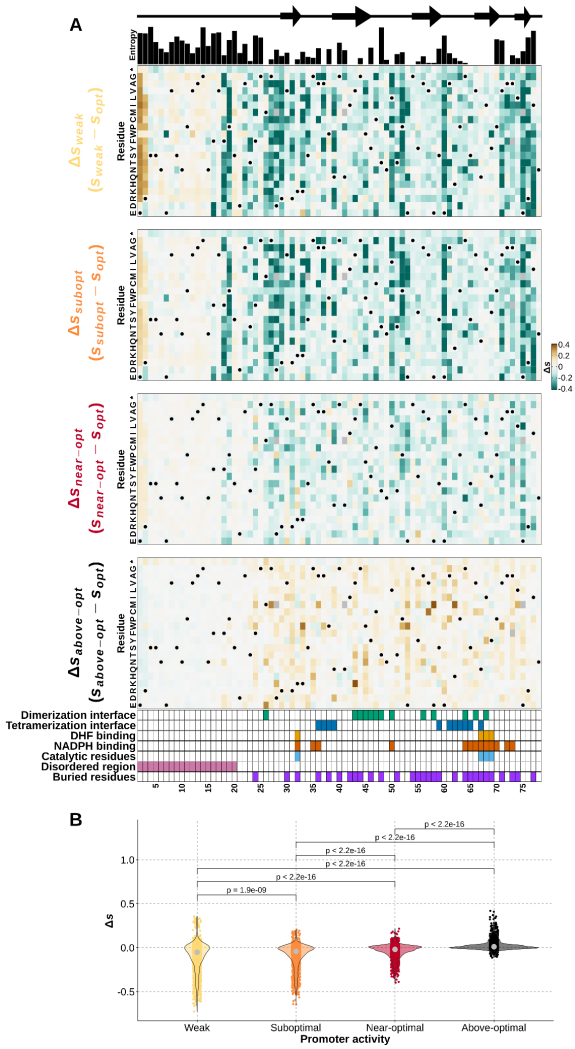


Fig. S12. Promoter activity-dependent differences in fitness effects become weaker as promoter activity level approaches the optimum. (**A**) Difference between the selection coefficients of an amino acid substitution at one of the non-optimal promoter activity levels and at the optimal promoter activity for the WT ($\Delta s_{non\text{-}opt} = s_{non\text{-}opt} - s_{opt}$). WT residues in panel A are labeled with dots. Entropy as a metric of variation in homologous sequences and secondary structures are shown at the top. Annotations at the bottom indicate residues that participate in the interfaces of the DfrB1 homotetramer, that come into contact with either the substrate (DHF) or the cofactor (NADPH), key catalytic residues, the disordered region, and buried sites. Interfaces are labeled as dimerization and tetramerization interfaces following (*22*). (B) Distributions of Δs for the different comparisons against s_{opt}. P-values were calculated using Wilcoxon's test for differences in means of paired samples.

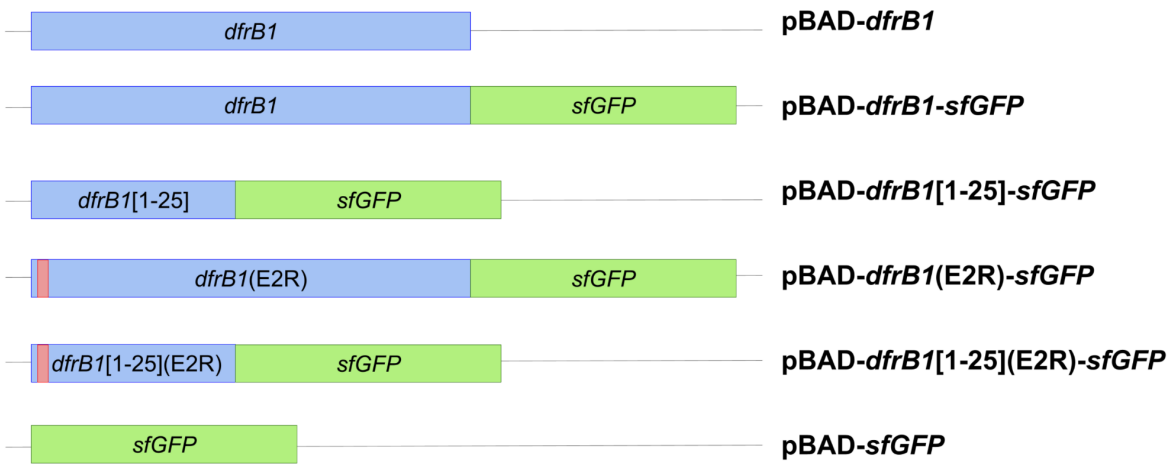


Fig. S13. Constructs used to evaluate the effect of E2R on protein abundance. The fluorescence of GFP fused to several constructs of DfrB1 was measured by flow cytometry (Fig. 4C). Constructs used include: 1) DfrB1 alone, 2) DfrB1 fused to GFP, 3) truncated DfrB1 (positions 1-25) fused to DfrB1, 4) E2R DfrB1 mutant fused to GFP, 5) E2R truncated DfrB1 mutant (positions 1-25) fused to GFP, 6) GFP alone.

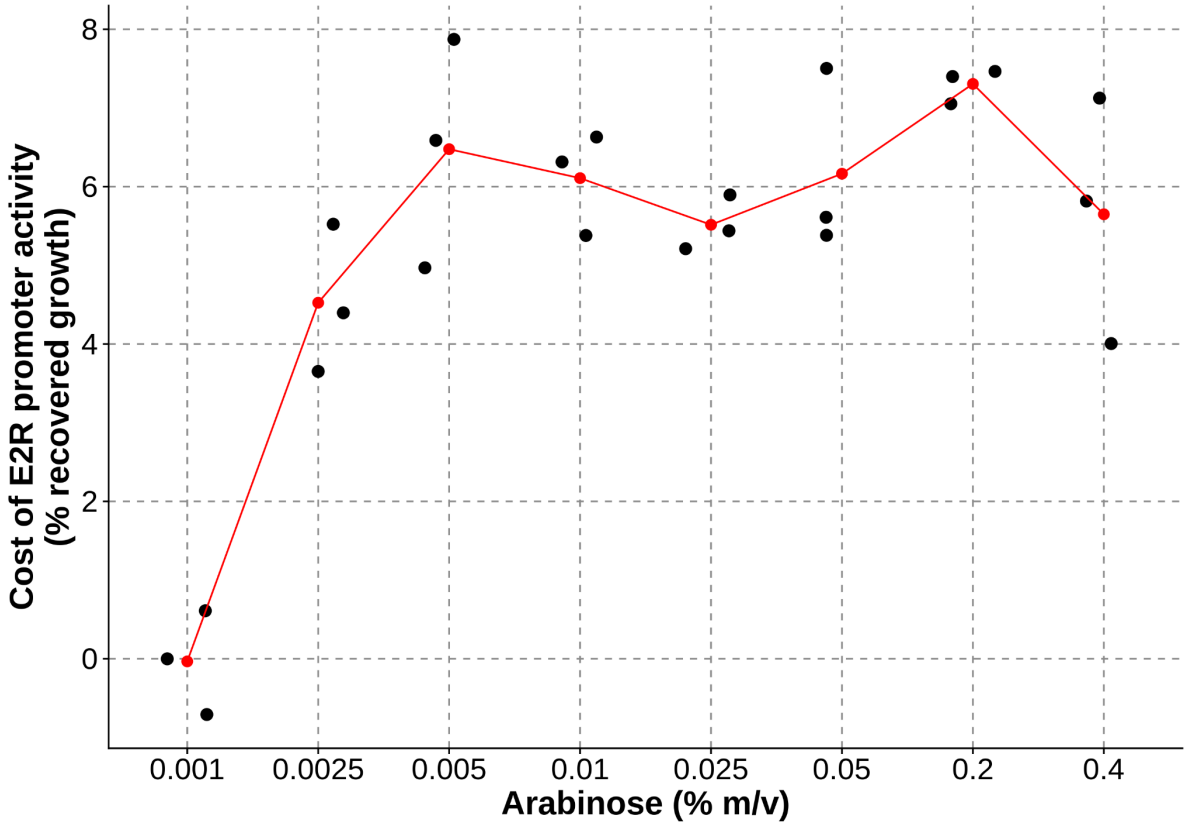


Fig. S14. Cost of promoter activity for WT DfrB1 and E2R mutant at different promoter activity levels. Cost of promoter activity for the E2R mutant in terms of percentage of recovered growth compared to the median recovery at 0.001 % arabinose, calculated using the same data from Fig. 4D. The red dots indicate the mean of three replicates for each arabinose concentration.

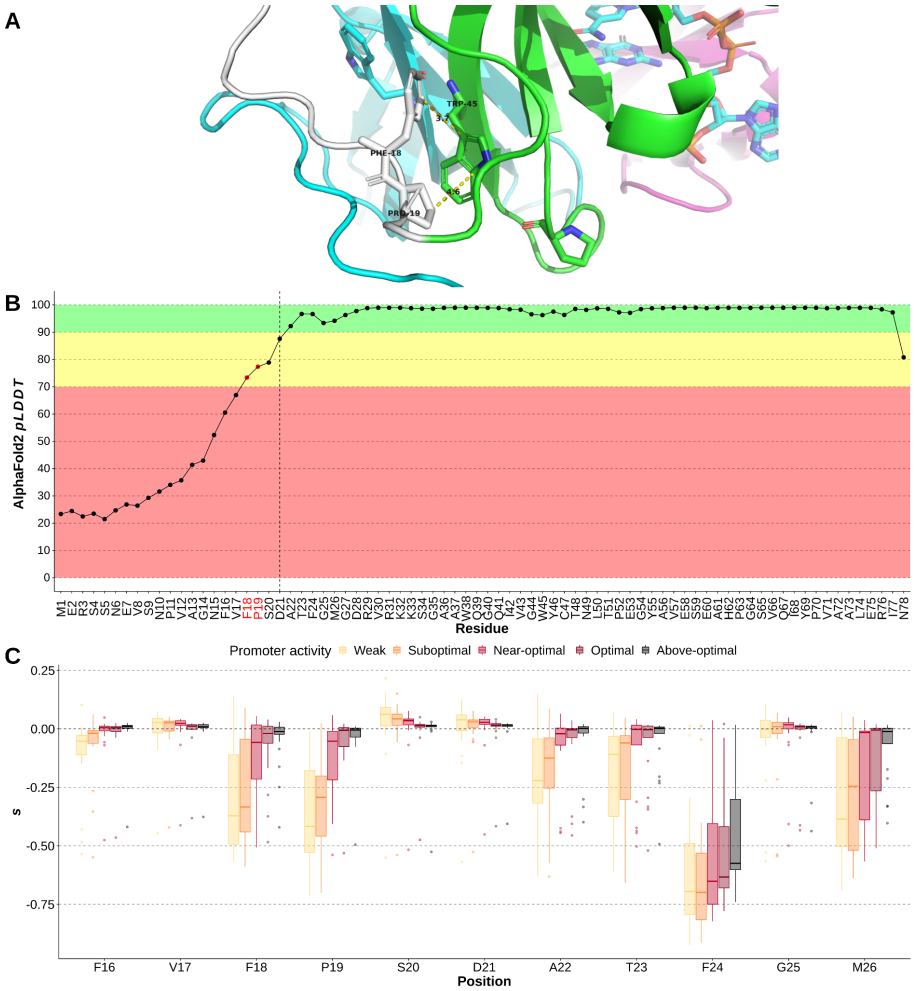


Fig. S15. Predicted interactions of W45 with F18 and P19. (**A**) Residues F18 and P19 are predicted to interact with W45 in an AlphaFold2 predicted model. Residues F18, P19, and W45, as well as measurements are labeled. The disordered region of subunit A is shown in white, the rest of subunit A is shown in green, subunit B is shown in magenta, and subunit C is shown in cyan. The structure was visualized with PyMOL (*76*). (B) *pLDDT* values assigned by AlphaFold2 to each of the predicted residues. Background colors indicate confidence levels: high confidence (*pLDDT* > 90, green), overall good backbone prediction (70 > *pLDDT* > 90, yellow), and low confidence (*pLDDT* < 70, red). Positions to the left of the dashed vertical line are not present in the crystal structure for DfrB1 (PDB: 2RK1). (C) Distributions of mutational effects for residues between positions 16 - 26 show that mutations at positions 18 and 19 are deleterious at low promoter activity levels but can be masked at the optimal promoter activity for the WT.

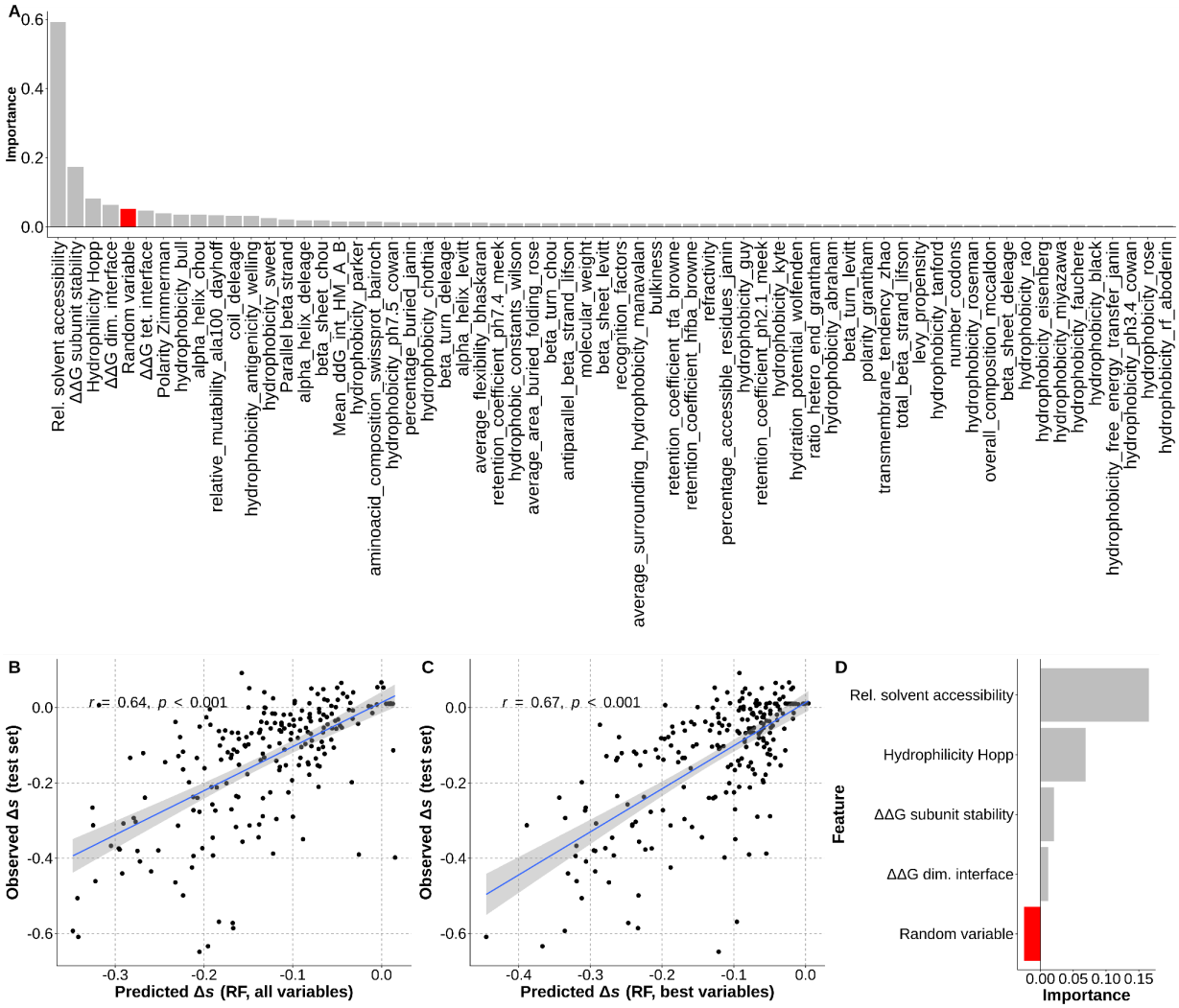


Fig. S16. Identifying the most relevant features used to predict promoter activity-dependent differences in selection coefficients. (**A**) Relative importance of all the features considered in the random forest model as calculated by the decrease in out-of-bag R^2 when permuting the values of that feature. A random variable was included to identify the features that contribute the most to the model. Table S9 contains a list of all variables used in this analysis. (B) Comparison of observed Δs_{weak} (s_{weak} - s_{opt}) versus predicted Δs_{weak} using a random forest (RF) regressor with all the features for the test set (20 % of the total dataset). (C) Comparison of observed Δs_{weak} (s_{weak} - s_{opt}) versus predicted Δs_{weak} using an RF regressor with all the top three features for the test set (20 % of the total dataset). (D) Relative importance of the top features in the final model as calculated by the decrease in out-of-bag R^2 when dropping one variable at a time and retraining the model.

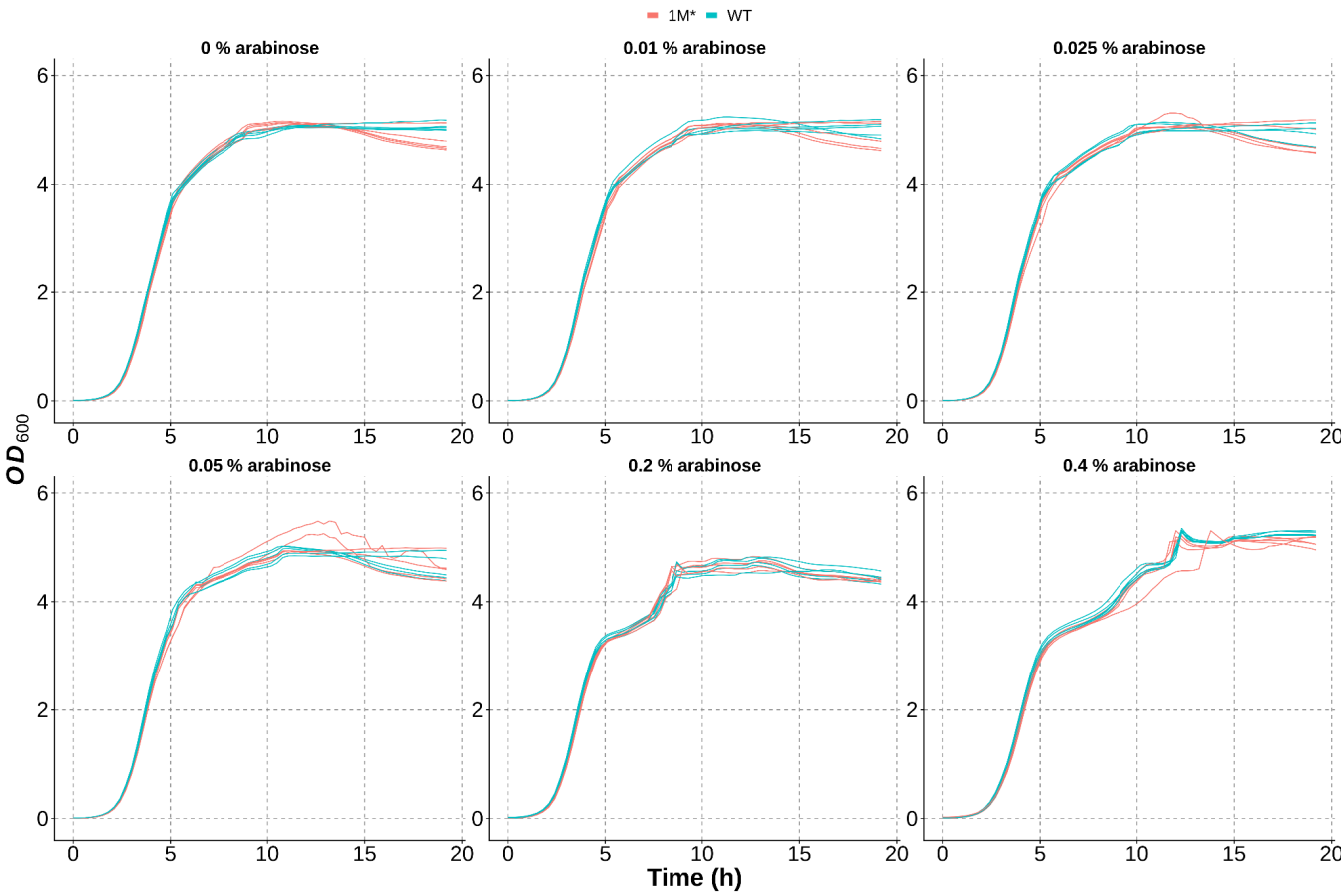
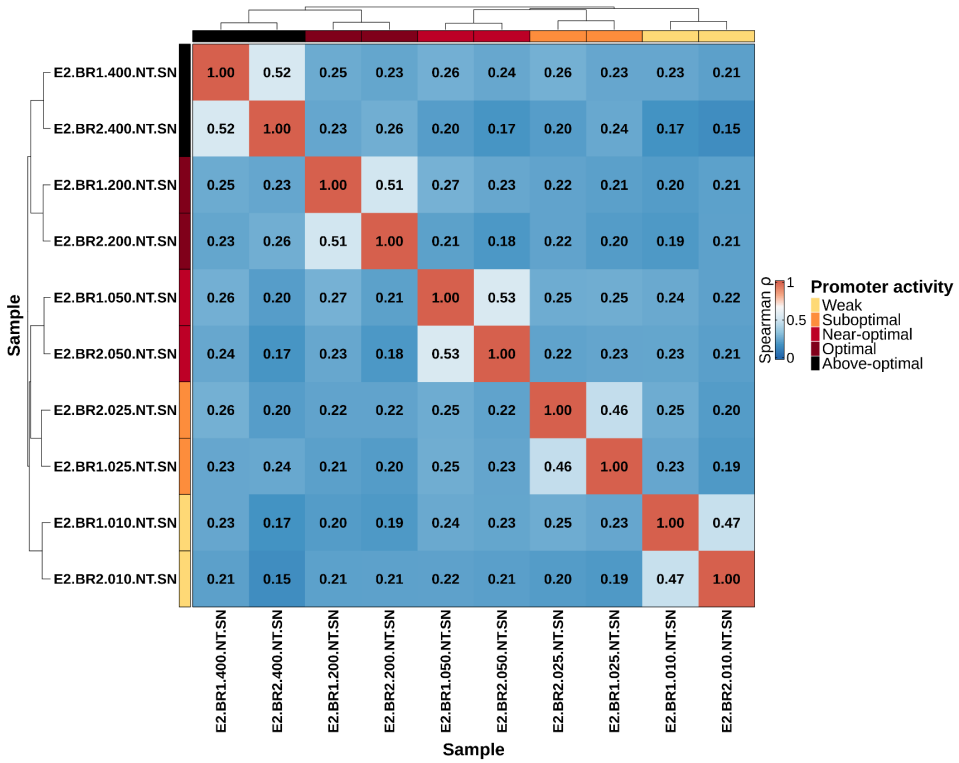


Fig. S17. No clear toxic effects of DfrB1 expression. Growth curves for cells transformed with WT DfrB1 or with a DfrB1 mutant with a stop codon at position 1 (AUG to TGA mutant - M1*) and grown in medium without TMP. Removal of the start codon should not prevent transcription but rather protein expression. A difference between the two alleles would therefore represent the cost of expressing the protein.



Experiment without TMP	ID	Biological replicates	Arabinose concentration (%)	Sequencing platform
2	E2.BR1.010.NT.SN	1	0.01	NovaSeq
	E2.BR2.010.NT.SN	2		
	E2.BR1.025.NT.SN	1	0.025	NovaSeq
	E2.BR2.025.NT.SN	2		
	E2.BR1.050.NT.SN	1	0.05	NovaSeq
	E2.BR2.050.NT.SN	2		
	E2.BR1.200.NT.SN	1	0.2	NovaSeq
	E2.BR2.200.NT.SN	2		
	E2.BR1.400.NT.SN	1	0.4	NovaSeq
	E2.BR2.400.NT.SN	2		

Fig. S18. Low signal-to-noise ratio in selection coefficients in the absence of selection for DfrB1 activity (no TMP). Spearman correlation between selection coefficients was estimated for different replicates of the experiment without TMP. Samples were named according to the table (bottom) extracted from Table S1.

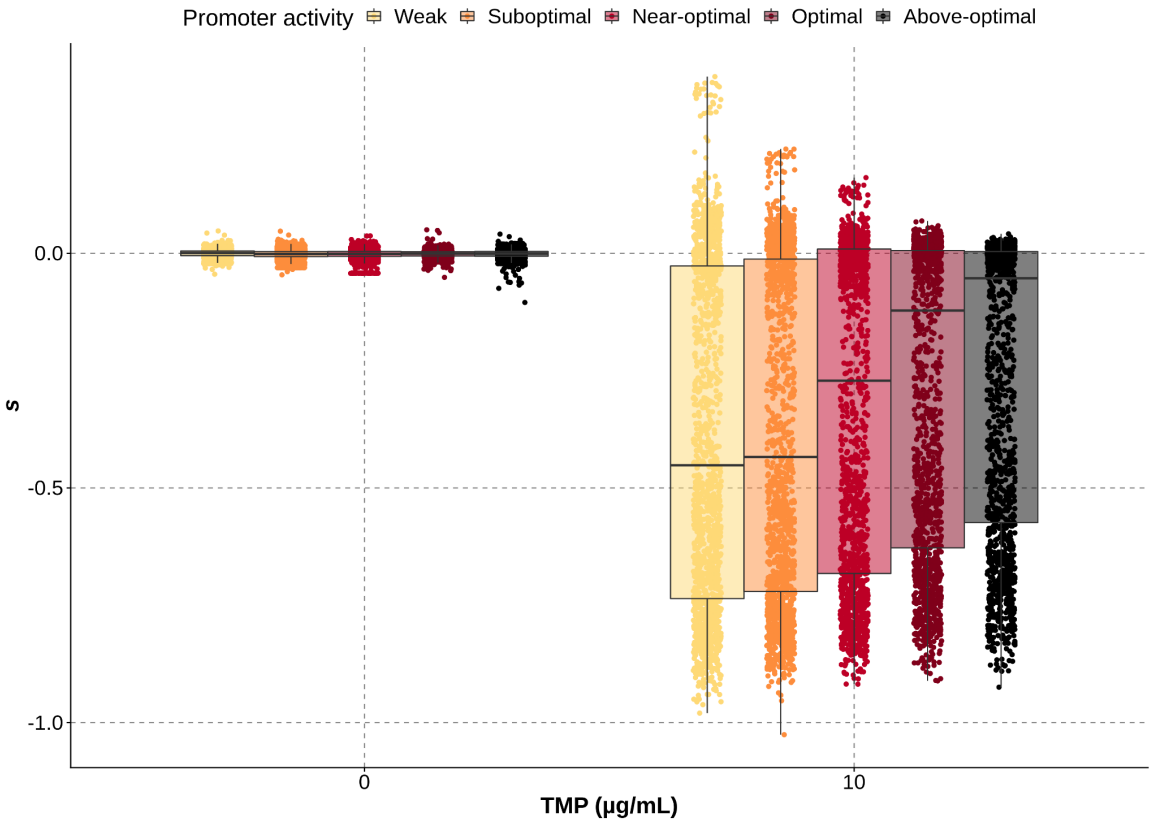


Fig. S19. Overall distributions of selection coefficients in the experiments with and without selection for DfrB1 activity (with and without TMP). The range of selection coefficients observed in the experiment with selection for DfrB1 activity is much broader than in the experiment without selection.

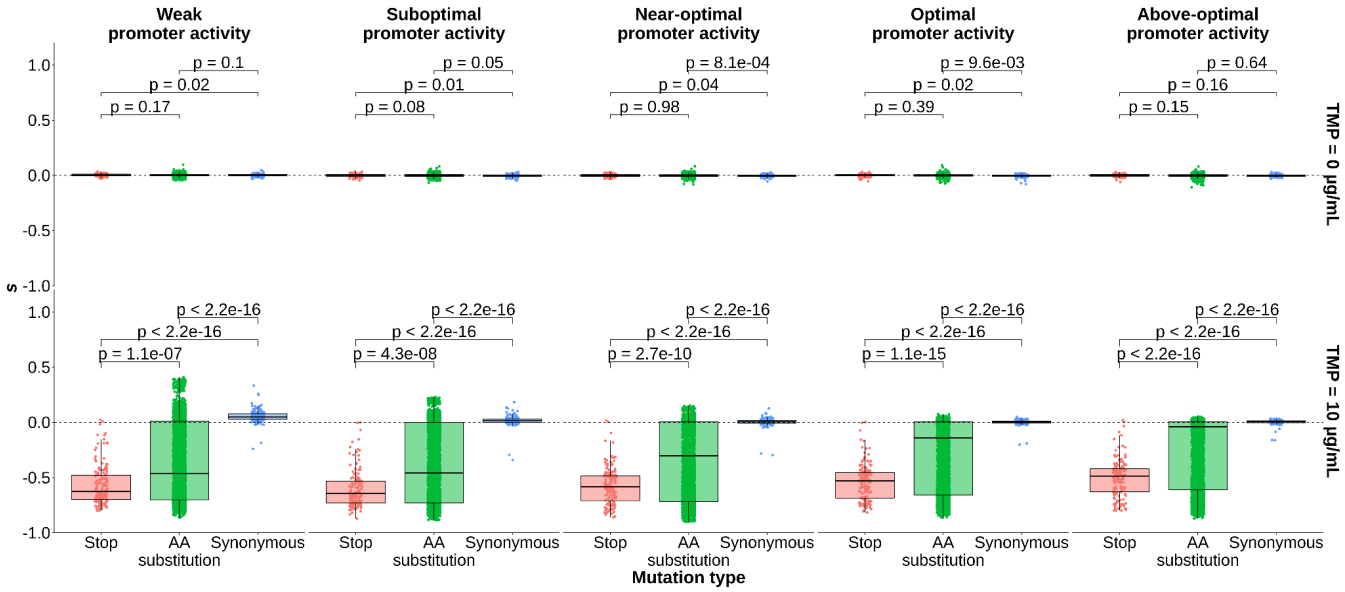


Fig. S20. Distributions of selection coefficients for stop codons, amino acid substitutions, and synonymous mutants. Mutations introducing amino acid substitutions or stop codons are deleterious in the experiment with selection for DfrB1 activity (with TMP), but not in the experiment without selection for DfrB1 activity (without TMP). P-values were calculated using Wilcoxon tests.

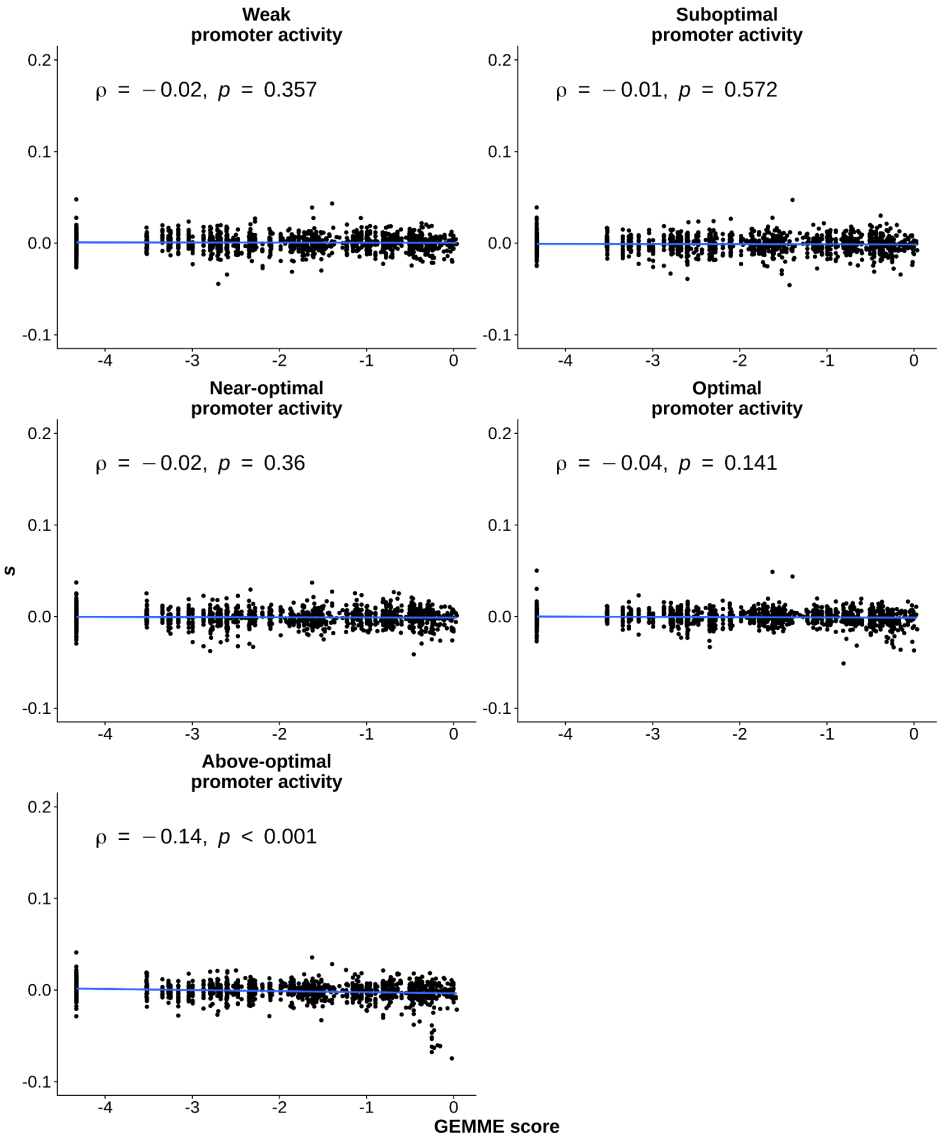


Fig. S21. Fitness effects measured in the DMS bulk competition experiment without TMP do not correlate well with fitness effects deduced variation observed in natural sequences. The GEMME scores indicating patterns of conservation and substitution (0 for residues observed as often as the WT DfrB1 residue at that position and negative for residues observed less often) are the same as in Fig. S11.

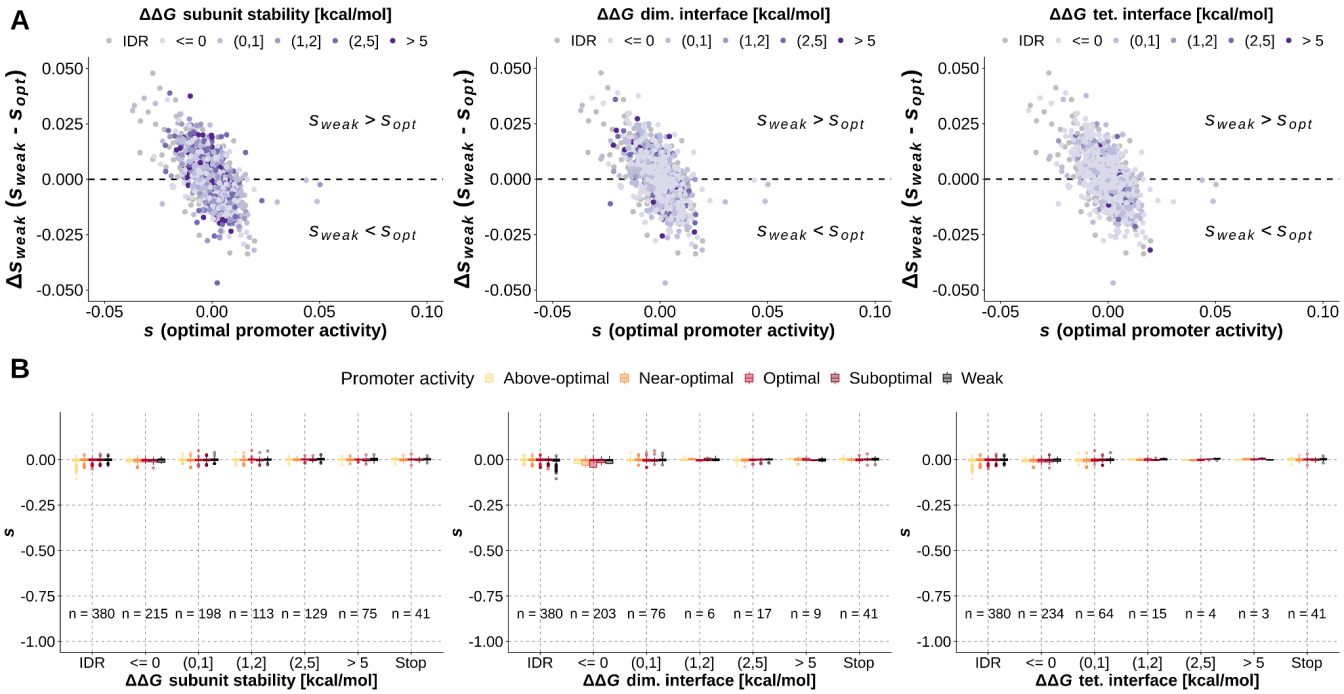


Fig. S22 Protein destabilization by itself does not result in deleterious effects in the absence of selection for DfrB1 activity. (**A**) Landscape of selection coefficients measured at optimal promoter activity (x-axis) and the promoter activity-dependent change in fitness effects (y-axis) labeled with respect to mutational effects on subunit stability and binding affinity. (B) Distributions of measured selection coefficients for mutants with different effects on protein stability or binding affinity at the interfaces. Stop codons from the functional protein core (residues 30 - 70) are shown for reference.