

Higher-order epistasis creates idiosyncrasy, confounding predictions in protein evolution

Karol Buda¹, Charlotte M. Miton¹ and Nobuhiko Tokuriki^{1*}

¹Michael Smith Laboratories, University of British Columbia, Vancouver - Canada

Corresponding author:

Nobuhiko Tokuriki

Michael Smith Laboratories, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

Tel: +1-604-822-8156

E-mail: tokuriki@msl.ubc.ca

Abstract

Epistasis shapes evolutionary outcomes during protein adaptation. In particular, when the effects of single mutations or mutational interactions are idiosyncratic, that is, unique to a genetic background, the predictability of protein evolution becomes greatly impaired. Here, we unveil a quantitative picture of the prevalence and role of idiosyncrasy in protein evolution by analysing 45 protein fitness landscapes, generated from seven enzymes. We found that mutational effects and epistasis are highly idiosyncratic across the landscapes. Idiosyncrasy obscured functional predictions of mutated proteins when using limited mutational data, and often continued to impair prediction upon incorporation of epistatic information. We show that idiosyncrasy stems from higher-order epistasis, and highlight examples where it permits, or restricts, evolutionary accessibility of certain genotypes. Our work suggests that idiosyncrasy deeply confounds predictions in protein evolution necessitating its incorporation into predictive models and in-depth exploration of its underlying molecular mechanisms.

Introduction

The evolution of new protein functions often requires the gradual accumulation of adaptive mutations. Accordingly, the fixation of multiple mutations along an evolutionary trajectory may cause a rewiring of intramolecular amino acid interaction networks that results in non-additive mutational effects, or epistasis¹⁻³. Epistasis causes a deviation from smooth, additive, and predictable mutational behaviours that – when rendered on a fitness landscape – creates a rugged mountain range with jagged peaks and valleys⁴⁻⁶. As evolution proceeds *via* the stepwise accumulation of single mutations, a rugged landscape topography dictates the ‘opening’ or ‘closing’ of accessible mutational paths across the landscape towards a given fitness peak^{7,8}. Thus, understanding, let alone predicting, protein evolution requires a robust description of epistasis and its ability to distort the sequence-function relationship⁴.

One approach to survey intramolecular epistasis in proteins relies on the functional characterisation of genotypes encompassing all possible combinations of a mutational subset that collectively alter protein function⁹⁻¹¹. Such investigations have demonstrated that there are a limited number of accessible mutational paths between the starting and endpoint genotypes, highlighting the evolutionary constraints exerted by epistasis¹¹⁻¹⁷. Recently, novel statistical approaches have been developed to obtain a more quantitative picture of epistasis, including higher-order epistasis (the non-additive effects of three or more mutations), embedded in these landscapes¹²⁻¹⁹. These statistical analyses typically provide a ‘global’ view of epistasis, where the averages of mutational and epistatic effects are captured across the combinatorial space^{12,15,17}. Such global approaches have demonstrated that the overall evolutionary changes within a combinatorial fitness landscape can be statistically recapitulated using the average effects of all individual mutations, supplemented only by the epistasis between pairwise and triplet mutations^{15,17}.

The global view of fitness landscapes can, however, have substantial shortcomings in scenarios where the effects of single mutations and epistasis are highly variable across different genotypes. For instance, if the effect of a mutation is idiosyncratic – *i.e.*, showing diverse functional contributions depending on the background genotype – its function in each genotype will significantly deviate from its global effect across the entire landscape^{20,21}. Since the majority of available mutational datasets characterise mutational effects in only a few backgrounds (generally single and double mutants in the protein’s wild-type (WT) background) most predictions made in protein evolutionary studies and engineering campaigns have the propensity to be heavily skewed by idiosyncrasy. Indeed, if the accessibility of a genotype is opened up or closed off by a strong, highly idiosyncratic, mutational effect, which is only present in a single (or very few) genotype(s) in the landscape, this idiosyncrasy will simultaneously dictate the evolutionary path followed by the protein whilst being lost in the global view, as it cannot be attributed to any epistasis captured by global effects. Hence, ubiquitous idiosyncrasy has the potential to confound our predictions and our understanding of evolutionary outcomes. However, we have little knowledge regarding the prevalence and roles of idiosyncrasy in protein fitness landscapes. If prevalent, what causes idiosyncrasy? To what extent does idiosyncrasy impact predictions in protein evolution?

In this study, we address these questions by systematically characterising 45 functionally annotated, combinatorially complete (for a subset of mutations), mutational landscapes. We explored the levels of idiosyncrasy across protein fitness landscapes and explore its impact on the accessibility of adaptive trajectories (Fig. 1). We first inferred idiosyncrasy from the extent of heterogeneity in single mutational effects and epistasis in these datasets. We then quantified the levels of WT idiosyncrasy by comparing the magnitude of every single mutation and epistatic contribution in the WT genotype to their

average across the landscape. Armed with a global and local statistical model, we unveiled the impact of idiosyncrasy on our ability to predict evolutionary- intermediates and endpoints across various fitness landscapes. Finally, we examined how idiosyncrasy affects the accessibility of a given evolutionary trajectory by permitting or restricting certain paths, and how these patterns translate to the molecular level.

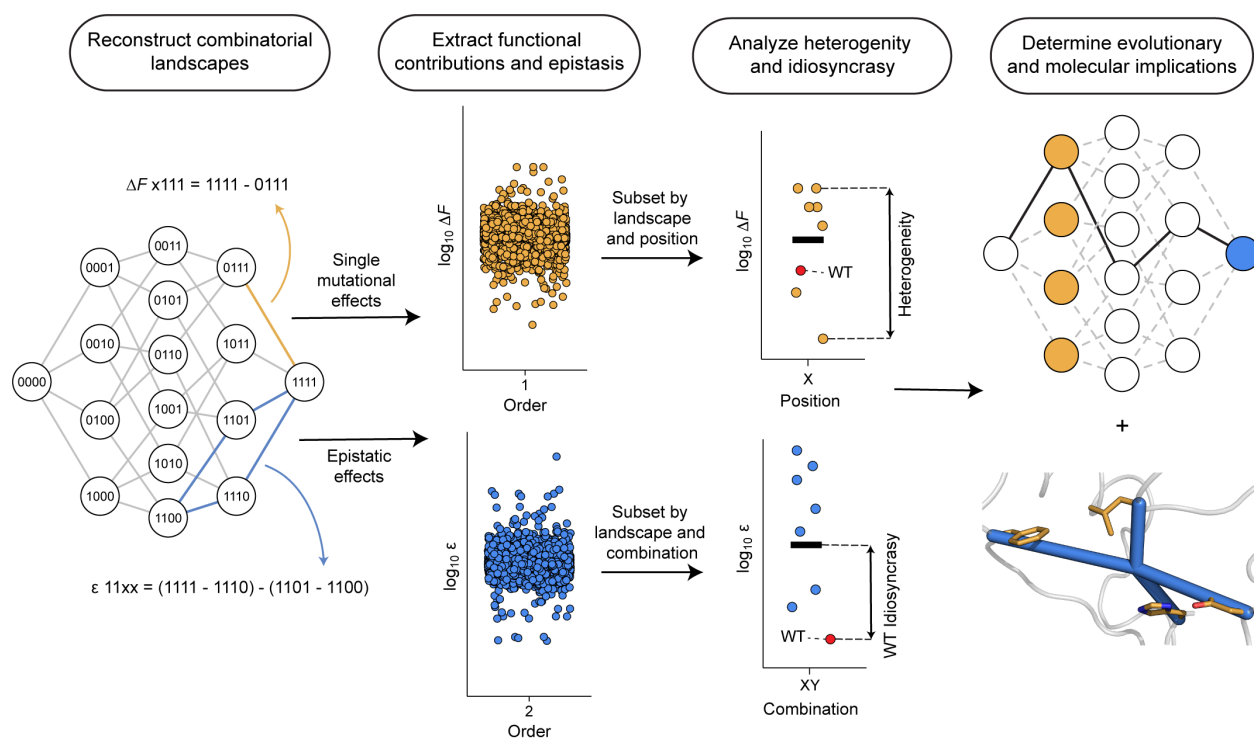


Fig. 1 | Analytical pipeline for characterising idiosyncrasy in protein combinatorial landscapes. From left to right, the data from empirical studies were processed to reconstruct combinatorial landscapes; in genotypes, 0 and 1 represent ancestral (WT) and derived (mutated) positions, respectively. We extracted the functional effect of mutations (ΔF) for each position (< 2 mutations), and epistasis (ϵ) for each combination (≥ 2 mutations). Heterogeneity and WT idiosyncrasy for single mutational- and epistatic-effects were captured using a statistical metric of spread and WT deviation from the mean (black bars), respectively. The impact of idiosyncrasy on evolutionary trajectories was analyzed and translated to the structural and molecular levels, e.g., single mutational (yellow) and epistatic (blue) data can be visualized on combinatorial landscapes and a protein structure to shed light on their evolutionary relevance.

Results

Statistical characterisation of 45 combinatorial landscapes. We collected several experimental studies that characterised the changes in protein function for a set of combinatorial mutations along an adaptive trajectory. The analyzed fitness landscapes – herein referred to as combinatorial landscapes – were restricted to datasets probing single mutations per position. These studies were filtered to ensure that the landscapes explored four or more positions ($n \geq 4$) and functionally characterised all possible combinations of these mutations (2^n variants). Of those studies, we only retained those reporting continuous data. Using these cut-offs, we obtained a working set of ten studies exploring seven different enzymes (Table 1). For some studies, the mutations were accumulated through directed evolution or protein engineering of a novel function (phosphotriesterase, PTE; β -lactamase, OXA-48; nitroreductase, NfsA)^{13,22–24}. For the remaining studies, the positions were identified from naturally occurring evolutionary trajectories, either through a retrospectively identified path using ancestral sequence reconstruction

(methyl-parathion hydrolase, MPH)^{14,25}, the presence of clinically relevant mutations (dihydrofolate reductase, DHFR and β -lactamase, TEM-1)^{11,26–28}, or in the case of alkaline phosphatase (AP), by using previously characterised active site residue mutations²⁹. The final dataset consisted of 56 unique mutations, of which 54 are located within the protein open reading frame, and two are located in a promoter region (in DHFR and TEM-1)^{11,23}. These were analyzed in 45 separate combinatorial landscapes with a total of 1,504 genotype-phenotype data points, some of which explored a set of positions for the same enzyme using a different substrate, inhibitor, or metal cofactor (Table 1 and Supplementary File 1).

Table 1 | Combinatorial landscapes analyzed in this study

Enzyme	No. of mutations	Conditions	No. of total measurements	Measured trait	Reference
OXA-48 ^a	4, 6, 6	2, 2, 2	288	IC_{50}	22
TEM-1 ^b	5, 4	1, 15	272	MIC, Growth rate	11,28
AP	5	1	32	k_{cat}/K_M	29
NfsA ^c	7, 7	1, 1	256	EC_{50}	24
DHFR ^d	4, 6, 5, 5	5, 1, 2, 2	272	IC_{50} , IC_{75} , k_{cat}/K_M and K_i	23,26,27
MPH ^e	5	8	256	Lysate activity	25
PTE ^f	6	2	128	Lysate activity	13

^a Three mutationally unique trajectories, where each was probed using two inhibitors

^b Ref. ²⁸ explored 15 inhibitors for the same set of four mutations

^c Two separate mutational trajectories for the same protein and substrate

^d Ref. ²⁶ explored four mutations using five different substrates; ref. ²⁷ explored both k_{cat}/K_M and K_i for two mutational trajectories

^e Ref. ²⁵ explored the same five mutations under eight different metal environments

^f Six mutations were explored using two substrates, one in ref. ¹³ and one outlined in this study (see **Methods**)

Next, the mutational data were processed to allow for a streamlined analysis using our computational pipeline (see Code Availability). Trajectories that explored different mutational combinations for the same enzyme were treated as separate combinatorial landscapes, as were the combinatorial landscapes characterising the function of the same subset of mutants across different substrates, ligands, or metals. Due to the variety of measured functions, ranging from direct physicochemical properties of the enzymes to indirect effects on the cellular phenotype, all enzyme functions were normalized relative to their WT background, providing us with the fold-change in enzyme function (F), which was then \log_{10} transformed (see Methods). Not accounting for the WT genotypes where $F = 0$, we obtained 1,459 F values for further analysis (Supplementary Data 1).

Idiosyncrasy in single mutational effects. To paint a comprehensive picture of the prevalence of epistasis in the selected proteins, we first extracted the functional effect of every single mutation at a given position across all available genetic backgrounds. The data were processed to provide the change in function (ΔF), or single mutational effect, of a given mutation across every genotype (see Methods). In a combinatorially complete fitness landscape of n mutations, a particular mutation occurs in 2^{n-1} distinct genetic backgrounds, hence, each combinatorial landscape contains $n \times 2^{n-1}$ ΔF values. Using this approach, we collected the 4,064 ΔF (Fig. 2a) and faceted them by the 214 unique mutation-substrate/inhibitor/metal pairs, simply referred to as ‘positions’, *e.g.*, the effect of a mutation at the same amino acid position for two different substrates is treated as a different ‘position’ (Supplementary Data 2). We chose to use a significance threshold of 1.5-fold for all analyses; this was the median error rate (calculated as $2 \times$ standard deviation, or 2σ) for all replicate measurements available in our data. Using this threshold, we found that the signs of the ΔF values were 21% negative, 36% neutral, and 43%

positive across all genotypes. This constitutes a relatively even split for the sign of the single mutational effects, despite a slight bias toward a positive effect, consistent with the fact that most mutations were originally identified due to their beneficial effect on function (Fig. 2a).

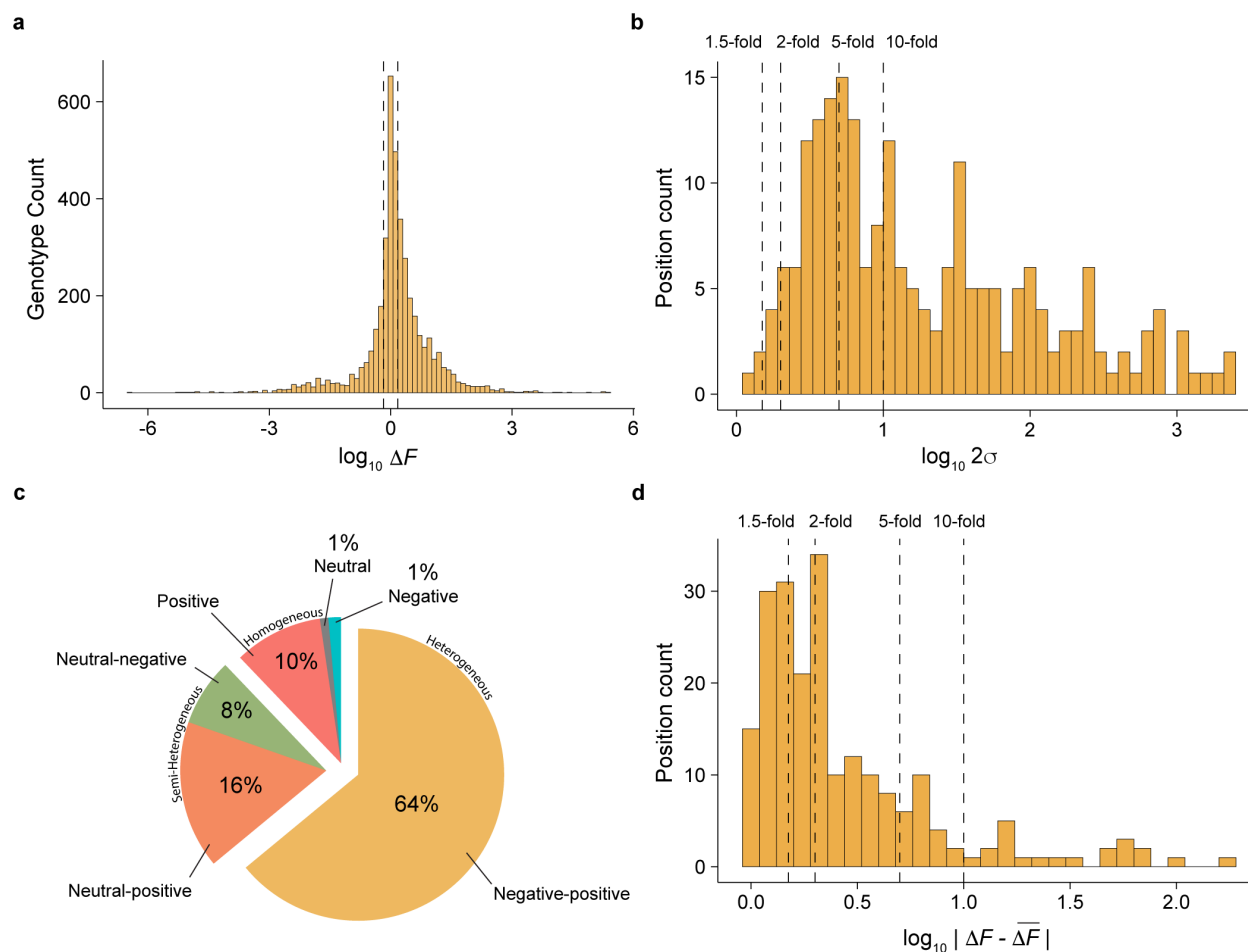


Fig. 2 | Descriptive statistics of single mutational effects. **a**, Distribution of single mutational effects ($\log_{10} \Delta F$) at every position across all genotypes for every combinatorial landscape. Dashed lines represent the 1.5-fold significance threshold that distinguishes negative, neutral, and positive effects, respectively. **b**, Distribution of the heterogeneity of the single mutational effects ($\log_{10} 2\sigma$ for the distribution of ΔF s) at each of the 214 positions. Dashed lines represent 1.5-, 2-, 5-, and 10-fold significance thresholds, respectively. **c**, Categorisation of the sign-changing behavior in ΔF distribution for each of the 214 mutations. **d**, Distribution of the WT idiosyncrasy quantified by the deviation of each of the 214 ΔF_{wt} from $\overline{\Delta F}$ at each position. Dashed lines representing 1.5-, 2-, 5-, and 10-fold significance thresholds, respectively.

Next, we characterised the variability in single mutational effects for a given position, herein referred to as heterogeneity, by plotting and analysing the spread in the distribution of ΔF at each position (Supplementary File 2). The heterogeneity in ΔF across different genetic backgrounds demonstrates the presence of extensive idiosyncrasy; accordingly, the degree of ΔF scatter for a position across various genotypes reflects the existence of epistasis of pairwise, and likely higher-order, interactions. To capture this spread, we computed 2σ for the ΔF at each position, which should encompass $\sim 95\%$ of the heterogeneity and inform on the spread around the mean single mutational effect for a given position. We deemed positions to be significantly heterogeneous when their $2\sigma > 1.5$ -fold. We found that heterogeneity prevails in our dataset: 99% (211/214) of positions exhibit significant

heterogeneity (Fig. 2b and Extended Data Table 2). Many positions exhibit much higher heterogeneity than the threshold: 70% (150/214) show a $2\sigma > 5$ -fold, and 53% (114/214) show a $2\sigma > 10$ -fold (Fig. 2b and Extended Data Table 2). We also characterised the extent of sign heterogeneity, *i.e.*, whether the sign of the single mutational effect varies between positive, negative, and/or neutral in different genotypic backgrounds. Only 12% of the positions showed sign homogeneity (retaining the same sign in all genotypes) and the remaining positions showed at least moderate sign heterogeneity, with 8% of positions contributing either a neutral or negative effect and 16% of positions displaying either a neutral or positive effect. Interestingly, 64% of positions showed full sign heterogeneity, showing background-dependent variance between positive and negative effects (Fig. 2c and Extended Data Table 2). These observations suggest that single mutational effects in the combinatorial landscapes are highly idiosyncratic – likely a result of pervasive epistasis.

Highly heterogeneous ΔF also suggest that the effect of a mutation can be idiosyncratic in a particular background compared to its global effect in the combinatorial landscape. We computed the difference in magnitude and sign of ΔF for a given mutation in the WT background (ΔF_{wt}) *versus* the mean ΔF ($\overline{\Delta F}$) of that mutation across all representative genotypes to determine the levels of WT idiosyncrasy (Fig. 2d). For 68% of positions, ΔF_{wt} deviates from $\overline{\Delta F}$ by >1.5 -fold (Fig. 2d and Extended Data Table 3), while the sign of ΔF_{wt} remains similar to that of $\overline{\Delta F}$ (only 3% of positions show a significant difference in sign). The idiosyncrasy in single mutational effects does not only apply to ΔF_{wt} , but also to the ΔF in other genotypes – 59% (2396/4064) of single mutants across every genotypic background also deviated >1.5 -fold from their $\overline{\Delta F}$ (Extended Data Table 4). This suggests that the ΔF_{wt} , or ΔF of any other genotypic background, is inaccurately captured by the global effect of that mutation within the combinatorial landscape. In other words, the magnitude of a previously observed single mutational effect is likely to vary in new genotypic backgrounds due to idiosyncrasy.

Idiosyncrasy in pairwise- and higher-order epistasis. The extensive idiosyncrasy in ΔF essentially stems from epistasis, therefore, we expanded our survey of idiosyncrasy in single mutational effects to interactions between mutations. Epistasis (ϵ), at any order, can be mathematically defined as the additional functional contribution originating from a new mutational interaction relative to all constituent lower-order interactions and single mutational effects. For example, our calculation of ϵ at the 3rd order (and higher orders) accounts for ϵ observed in the 2nd order, in contrast to the traditional calculation of the difference between observed function and predicted function using only the sum of the single mutational effects. Thus, we used the following equation to calculate 2nd order epistasis between mutations:

$$\epsilon_{p1p2} = F_{p1|p2} - F_{p1} - F_{p2} \quad (1)$$

Where $p1$ and $p2$ represent the single mutations at hypothetical positions ‘1’ and ‘2’, respectively, $p1|p2$ represents the combination of mutations $p1$ and $p2$, ϵ is the epistasis coefficient, and F is the normalized function. Higher-order epistasis was calculated using a similar approach as in equation 1 (see Methods). The extracted ϵ coefficients for all combinations across all landscapes were probed using the same metrics applied to the ΔF , to capture the extent of idiosyncrasy in epistasis.

Compared to the 214 surveyed positions, we extracted 419 pairwise-, 438 triple-, and 267 quadruple- combinations. Since a few landscapes only covered 4 mutations, 22/267 quadruple combinations could not be examined as they only occurred in one genetic background, leaving 245 quadruple combinations viable for analysis. Ultimately, we collected 8968 values for ϵ encompassing the

2nd to the 4th order (Supplementary Data 3). Surprisingly, we found that nearly all (99% (413/419)) pairwise interactions are significantly heterogeneous (Fig. 3a). This proportion decreased at less stringent thresholds, nonetheless, even with a 10-fold significance cut-off, 41% of pairwise interactions remained significantly heterogeneous (Extended Data Table 2). Interestingly, this trend remains consistent at higher orders of epistasis: 96% (420/438) of triple interactions, and 98% (240/245) of the quadruple interactions exhibit significant heterogeneity (Fig. 3a). The impact of high heterogeneity in epistasis was also seen in their sign variability – for pairwise combinations, 71% (295/419) show full sign heterogeneity (Fig. 3b). Only a handful of them (5%) exhibited sign homogeneity, and the remaining combinations were either neutral-negative (9%) or neutral-positive (15%). For higher-order combinations, sign heterogeneity remains high: 54% for 3rd order, and 52% for 4th order combinations (Fig. 3b and Extended Data Table 2).

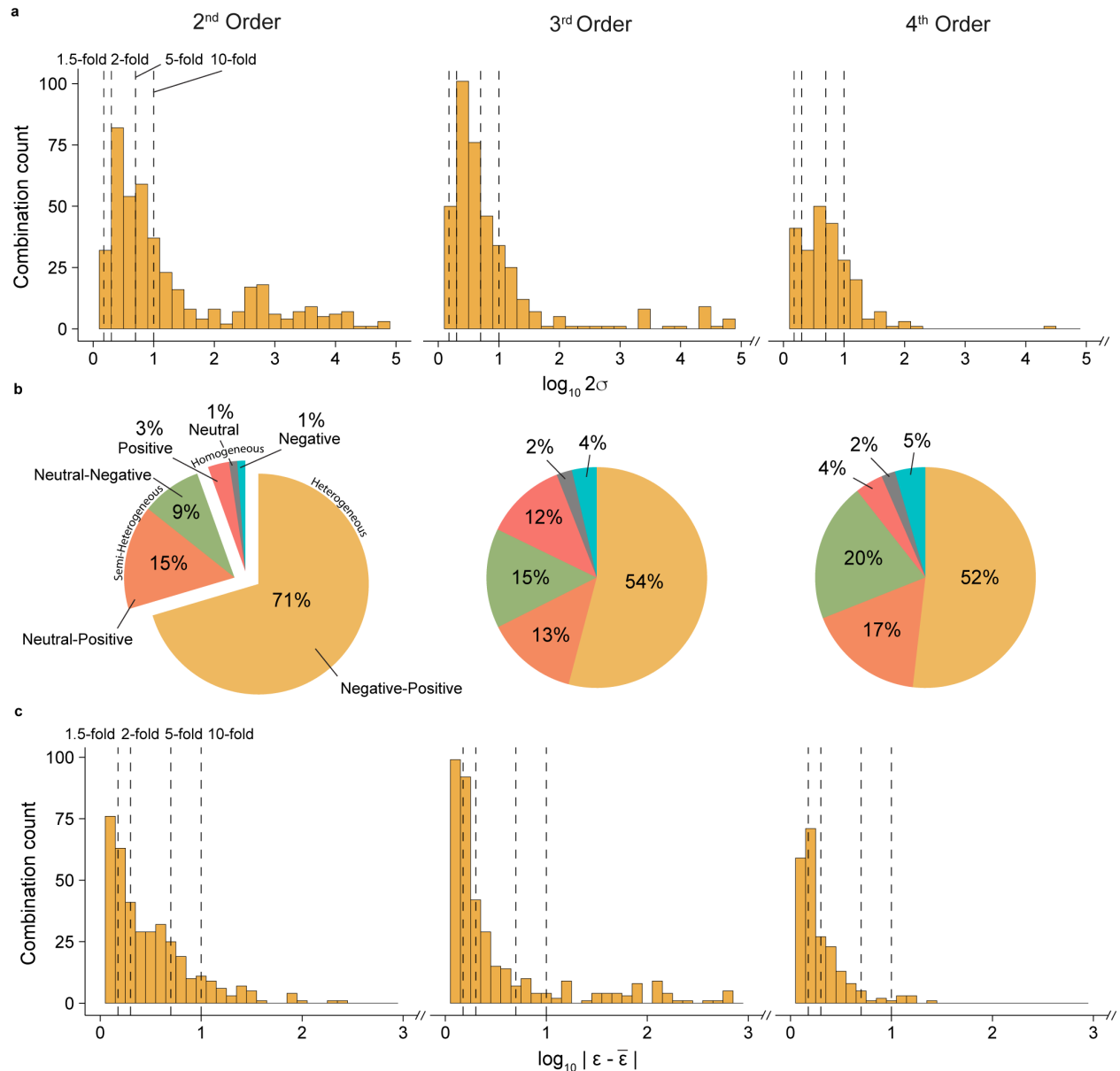


Fig. 3 | Descriptive statistics of epistasis. **a**, Histogram of $\log_{10} 2\sigma$ of ε at each order for every mutational combination, with annotated dashed lines representing 1.5-fold, 2-fold, 5-fold, and 10-fold significance thresholds, respectively. Data omitted due to x-axis scaling can be found in **Extended Data Fig. 1a**. **b**, Categorisation of the ε at each order for all mutational combinations by sign. **c**, The distribution of the WT idiosyncrasy in epistasis quantified *via* the absolute difference of ε_{wt} from $\bar{\varepsilon}$ at each combination. Dashed lines represent 1.5-, 2-, 5-, and 10-fold significance thresholds, respectively. Data omitted due to x-axis scaling can be found in **Extended Data Fig. 1b**.

Like single mutational effects, the WT background epistasis (ε_{wt}) showed a high deviation from the mean ($\bar{\varepsilon}$) for each combination. More than half (64% (270/419)) of pairwise, (56% (244/438)) of triplet, and (62% (153/245)) of quadruplet ε_{wt} deviated from $\bar{\varepsilon}$ by more than 1.5-fold (Fig. 3c). As with the ΔF , this high deviation was also representative of other non-WT backgrounds (Extended Data Table 4). Interestingly, more combinations showed significant sign discrepancy in ε_{wt} *versus* $\bar{\varepsilon}$ than those for ΔF . Although still in the minority, 13% (53/419) of pairwise-, 5% (34/438) of triplet-, and 8% (12/245) of quadruplet- interactions had a significantly different sign effect in the ε_{wt} than the $\bar{\varepsilon}$. These findings

suggest that epistasis itself remains idiosyncratic, even at increasing orders. The persistence of idiosyncrasy at the 3rd and 4th orders of epistasis suggests that epistasis at the 5th (and likely higher) order(s) continues to impose a genotypic context dependence even on triplet and quadruplet mutational interactions.

Idiosyncrasy confounds predictability. Given the strong prevalence of idiosyncrasy in single mutational effects and epistasis, we asked to what extent does idiosyncrasy impair the predictability of genotypic functions in the landscape? We computed functional predictions using two different models and examined the discrepancies between them. The first model collected the ΔF_{wt} and ε_{wt} for every position and combination and computed the function at every genotype using the additive assumption of that genotype's constituent ΔF and ε values (Fig. 4a and Methods). We refer to this as the 'WT-background model', because the model relies on limited data from the WT background, namely ΔF_{wt} and ε_{wt} , as opposed to any other, or all, ΔF and ε across the landscape (Fig. 4b). This model is inspired by the local view of epistasis^{16,17}, which stems from the stepwise process employed in protein engineering and evolutionary studies. We compared the WT-background model to a global model that used $\overline{\Delta F}$ and $\bar{\varepsilon}$; by averaging ΔF and ε values across the landscape this model aims to minimize the idiosyncratic noise of the system. Our global model of choice was a linear regression with interaction terms, simply referred to as the global model, which is conventionally employed in combinatorial landscape analyses (Fig. 4a)^{11,13,14,16,17,25}. This comparison aims at exposing how a deviation between mutational- and epistatic- effects in WT background *versus* their corresponding mean effects can confound predictions of evolutionary outcomes.

We computed the predicted functions, using both the WT-background and global models, with gradually increasing orders up to the 4th order (Fig. 4a and 4b). To assess the quality of the functional prediction performed by each model, we measured the absolute error (AE) of the predicted *versus* observed function for the "endpoint", or most derived, genotype in each landscape (Fig. 4c and Methods).

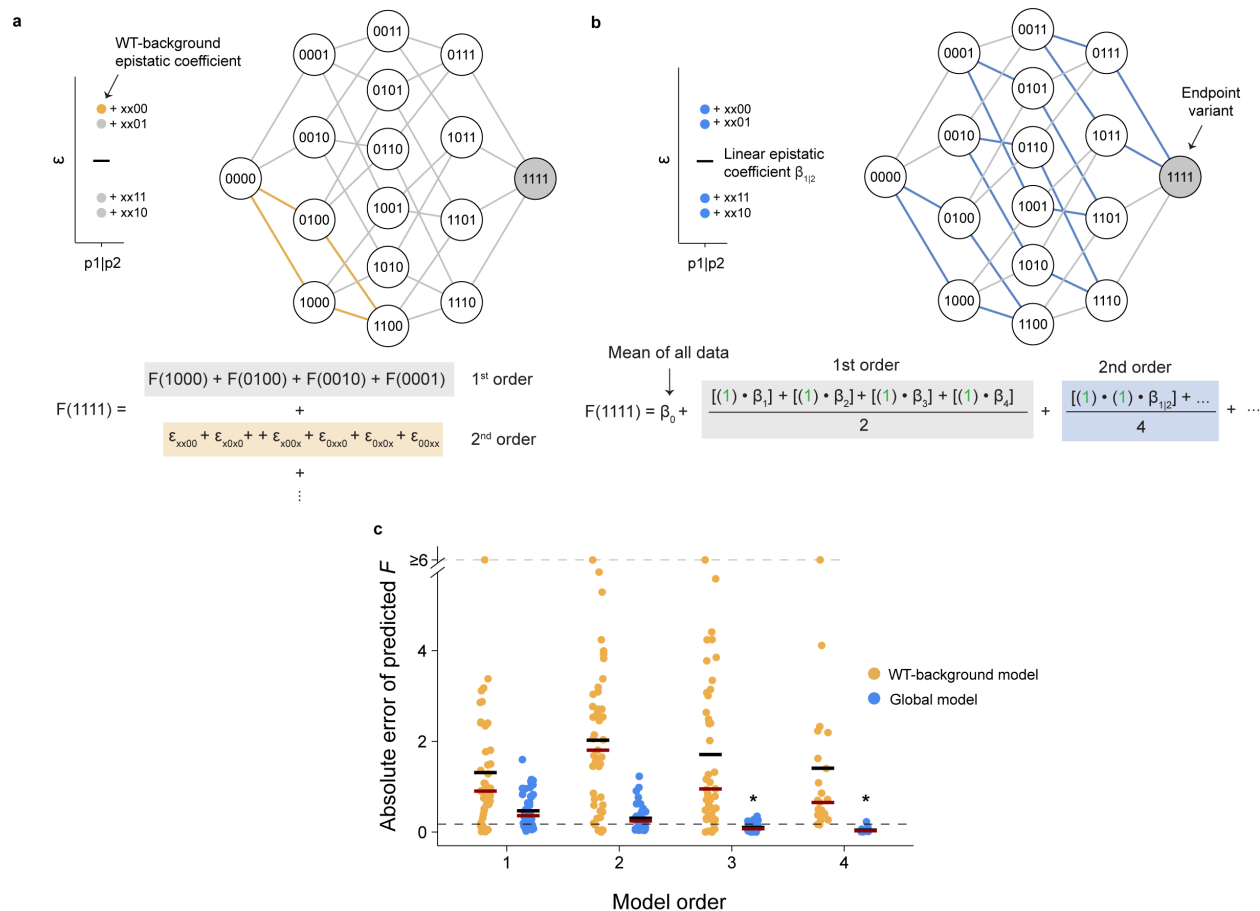


Fig. 4 | Differences in WT-background and global models. **a-b**, Predicting the function of an endpoint variant in a hypothetical landscape using the WT-background (**a**) and global (**b**) model. **a**, The WT-background only uses epistasis in the WT background to extract the epistatic coefficient (ϵ_{xx00}) and uses it to predict the function (equation below and **Methods**). **b**, The global model computes epistatic coefficients for an interaction, here for positions 1 and 2 (β_{12}), from the mean epistasis between mutations across all genotypes that contain this interaction (equation below and **Methods**). **c**, Absolute error of the predicted function (F) at orders 1, 2 and 3 ($n=45$), and at order 4 ($n=23$). The absolute error means and medians from the global (blue) and WT-background models (yellow) are shown as black and red bars, respectively. The 1.5-fold significance threshold is depicted as a dashed line, with an asterisk (*) marking the model and order with a significant mean (one-sided t-test; $p = 9.67 \times 10^{-7}$ and 1.83×10^{-12} for global model order 3 and 4, respectively).

Given the discrepancy in the amount of data accessed by the models, as expected, the mean AE of the global model outperformed the WT-background model at the lower orders. In the 1st order, 13% (6/45) and 29% (13/45) of variants were successfully predicted by the WT-background and global model, respectively. The global model's prediction accuracy increased with each incorporated order. Furthermore, the incorporation of 3rd order information was largely sufficient for predicting the function of most endpoint variants below the significance threshold ($p < 0.05$; Fig. 4b and Extended Data Table 5). This suggests that average effects, encompassing relatively minimal epistatic information, can recapitulate the topology of the combinatorial landscapes. By contrast, the WT-background model showed high AEs with a marginal improvement in median AE, but not mean AE, even upon introducing 4th order epistatic information (Fig. 4b and Extended Data Table 5). This poor performance of the WT-background model suggests that the ΔF_{wt} and ϵ_{wt} values were not sufficient predictors of the function of genotypes with multiple mutations, highlighting how idiosyncrasy confounds functional predictions in protein fitness landscapes.

Idiosyncrasy disrupts the interpretability of adaptive trajectories. Next, we performed several analyses to assess the extent to which idiosyncrasy impairs the predictability of adaptive trajectories. To this end, we focused on 11 adaptive landscapes, *i.e.*, landscapes where the substrate or ligand is assumed to be the primary selection pressure that led to the accumulation of the probed mutations. For each landscape, we retained a single most accessible path, in which the most functionally advantageous mutation is fixed at each step. We then computed a predicted function for each genotype along the most accessible path using ascending orders of the WT-background model (Fig. 5a and Extended Data Table 6). However, much as with the endpoint prediction, we again saw that for most genotypes the WT-background model failed to predict function accurately at all orders. Interestingly, some highly mutated genotypes along the trajectory were, in fact, better predicted by lower-order information. This indicates that when the WT-background model incorporates extensive idiosyncrasy at a given order, its predictive power is impaired relative to a model with less mutational information (Fig. 5a and Extended Data Table 6).

Next, we sought to gain a statistical picture elucidating how idiosyncrasy can open or close certain evolutionary trajectories in the adaptive landscapes. First, we defined the genotype $n+1$ as accessible if its function increases, or does not decrease by more than 1.5-fold, compared to the ancestral n genotype. In other words, we enabled nearly neutral steps to occur by considering that a 1.5-fold decrease in function is insignificant. Then, we examined how newly introduced interactions (*e.g.*, 4th order epistasis for a quadruple mutant, but not 3rd order or lower) affect the accessibility of a genotype, and marked the genotype as ‘affected by epistasis’ if removing the new interaction ($\epsilon_{\text{new}} = 0$) meant that the genotype’s accessibility was affected. Furthermore, a genotype was considered ‘affected by idiosyncratic epistasis’ if its accessibility was affected upon converting the effect of the new epistatic interaction to the average epistasis of that interaction ($\epsilon_{\text{new}} = \bar{\epsilon}$) across the landscape.

By looking at all genotypes across the adaptive landscapes, we found that new epistatic interactions affected the accessibility of 49% (288/583) of genotypes. Furthermore, 78% (211/288) of these genotypes showed different accessibility specifically due to idiosyncratic epistasis (Fig. 5b). Of course, many of these genotypes are unlikely to be accessed at all, as early mutations with greater functional contributions will be selected first, consequently affecting the chance of visiting many intermediate genotypes. Thus, we further refined our accessibility metrics by focusing on the genotypes in the most accessible path for each adaptive landscape. Of these most accessible genotypes, the accessibility of 26% (11/42) was significantly affected by new interactions, and 73% (8/11) of those genotypes were idiosyncratic. Hence, although in the most accessible trajectories new epistatic interactions are generally less likely to affect the accessibility of a genotype, when they do impact the evolutionary path, they are usually idiosyncratic.

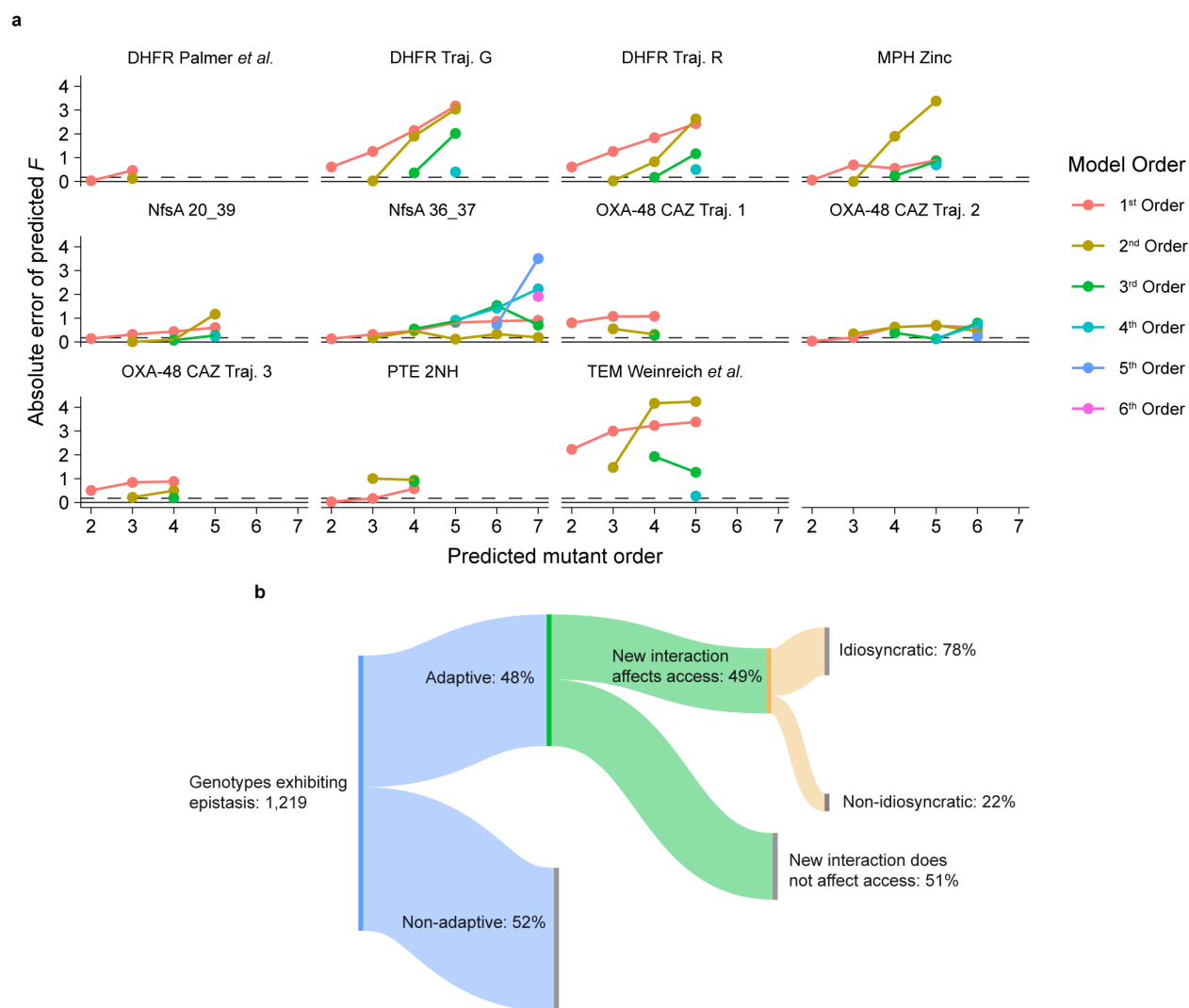


Fig. 5 | Impact of idiosyncratic epistasis on prediction and genotype accessibility. **a**, Absolute error of the functional prediction for each genotype along the most accessible path in 11 adaptive landscapes. Colors indicate the order at which the WT-background model was used. **b**, Sankey plot of epistasis and its link to genotype accessibility in adaptive landscapes. Epistasis encompasses all ϵ values of newly introduced interactions between ≥ 2 mutations that show > 1.5 -fold magnitude in effect. Genotypes with new significant epistasis were filtered by their occurrence in the 11 adaptive, or 34 non-adaptive landscapes, then by modulation of accessibility by new interactions, and finally by idiosyncrasy.

Through this analysis, we identified two trajectories exhibiting particularly interesting evolutionary dynamics caused by idiosyncrasy, and decided to probe their genetic and molecular underpinnings. The first trajectory was the most accessible path along the PTE evolution toward the arylesterase activity^{13,30–32}. Along this trajectory, the highest functional peak is reached at genotype D233E-H254R-L271F-F306I with $\sim 1,700$ -fold improvement in function relative to WT (Fig. 6a). The accessibility of this genotype is permitted by a strong quadruple epistatic interaction between these positions, resulting in a 7.4-fold increase in function relative to the contribution from all single mutational, pairwise- and triple- interaction effects (Fig. 6b). Without this 4th order epistasis, this genotype would become inaccessible, as it would lead to a ~ 4.5 -fold decrease in function relative to D233E-H254R-F306I. Interestingly, this genotype would also be inaccessible if the effect of the quadruple interaction was represented by its $\bar{\epsilon}$ (Fig. 6b). In fact, $\bar{\epsilon}$ captures the incompatibility of the quadruple epistasis with L272M

and I313F, illustrating the idiosyncratic nature of this quadruple interaction within the PTE landscape. In other words, this idiosyncratic quadruple interaction provides an evolutionary solution that permits access to the functional peak, but can quickly become disrupted by negative higher-order epistasis from other mutations. Idiosyncrasy can also, on the other hand, be restrictive. In the MPH landscape, five adaptive mutations were found to confer methyl-parathion hydrolase activity to MPH^{14,25}. Yet, a pair of individually beneficial mutations exhibit negative epistasis when combined: the L72R-F273L interaction restricts the transition from the single (F273L) to double mutant (L72R-F273L) due to a ~1.7-fold decrease in function (Fig. 6c). However, the mean pairwise epistasis between L72R and F273L is significantly positive (+4.0-fold), suggesting that the L72R-F273L pair becomes beneficial later in the trajectory. Indeed, this sign change in ϵ results from the early fixation of $\Delta 193S$, which greatly stabilizes the L72R-F273L interaction, and is further reinforced by positive higher-order interactions with H258L and I271T (Fig. 6d). Thus, higher-order epistasis appears to be the key mechanism underlying the compensation of this initially idiosyncratic antagonism, later fostering the pair's synergy. These two examples demonstrate how idiosyncrasy is a by-product of higher-order interactions – whether synergistic or antagonistic – and how it can create localized permissive or restrictive dynamics in combinatorial landscapes that are concealed by the reductionism of the global view.

Lastly, we explored whether the high levels of idiosyncrasy could be explained by other features in our adaptive trajectories. For example, it has been debated whether the functional contribution of a mutation is affected by the functional magnitude of the genotype in which it occurs³³. Indeed, several adaptive trajectories in our study exhibited diminishing returns topology, where the functional contribution of mutations decreased as they were introduced into fitter, and more derived, backgrounds (Extended Data Fig. 2a). This trend has often been attributed to negative epistasis arising between mutations that originally appeared beneficial, a phenomenon called diminishing return epistasis^{18,20,21,34}. However, upon probing ϵ in the adaptive landscapes, we found no trend in ϵ based on the functional level of intermediates, suggesting no evidence of the dominance of diminishing return epistasis (Extended Data Fig. 2b). As such, the patterns of diminishing returns observed in the adaptive landscapes are likely the result of other mechanisms, such as beneficial mutational exhaustion, whereby the most beneficial adaptive mutations are fixed in early rounds of evolution, quickly depleting the pool of beneficial mutations as evolution progresses.

We also explored idiosyncrasy from the context of protein structures. Since intramolecular protein epistasis is attributed to the interaction between mutations, we analyzed the spatial proximity between the epistatic positions as a function of the magnitude of epistasis between said positions. However, we observed no significant correlation between the distance connecting the α -carbons of two residues and the absolute magnitude of their pairwise ϵ (Extended Data Fig. 3). Thus, other factors have to be considered when characterising the relationship between spatial proximity of mutations and epistasis, for instance, non-specific epistatic effects³⁵, side chain interactions, or perhaps complex mutational interactions including conformational dynamics or ligand and substrate relays¹.

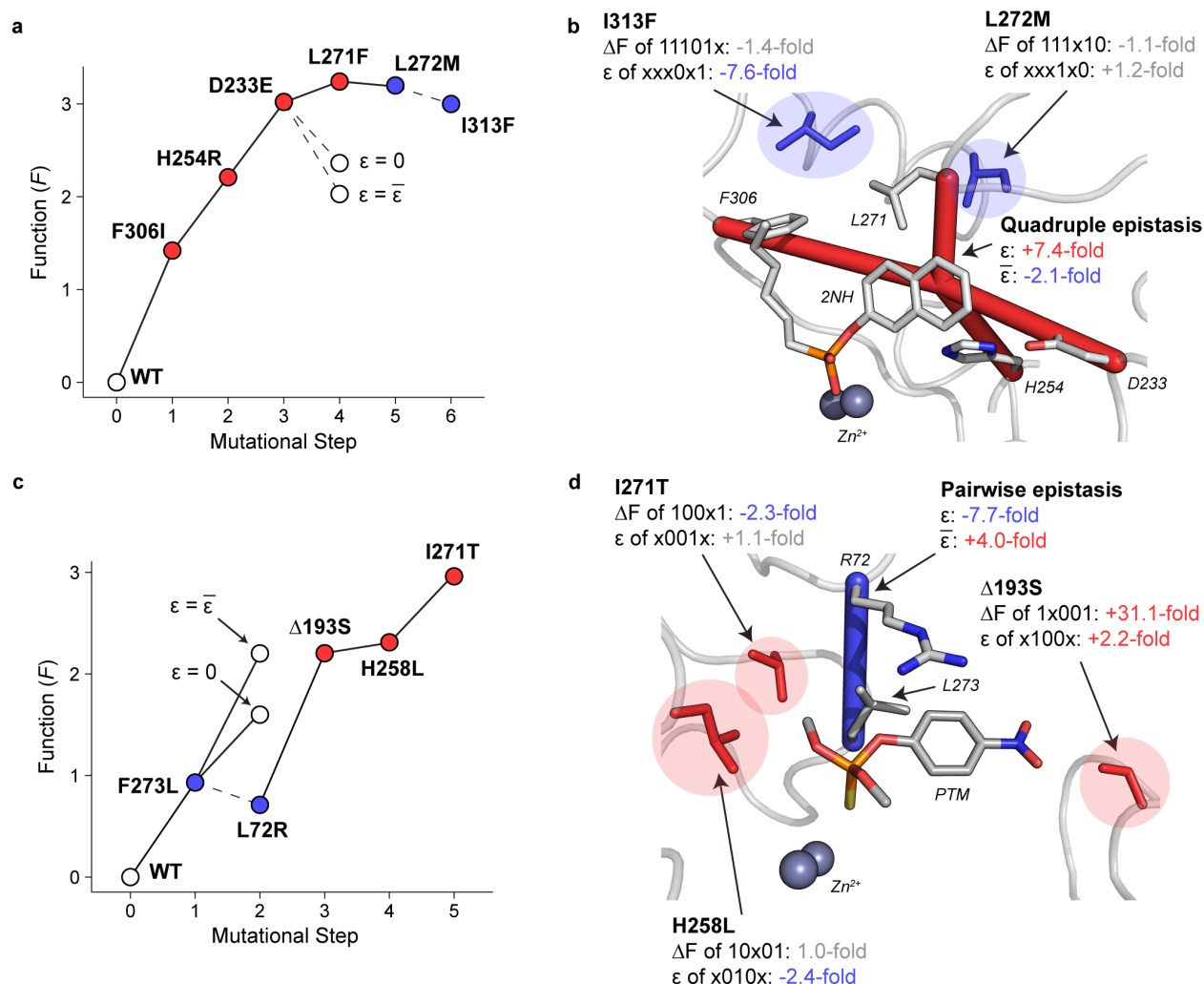


Fig. 6 | Structural representation of idiosyncrasy and its effect on evolutionary trajectory accessibility. **a, c**, Function of genotypes along **a**, the PTE trajectory toward arylester (2NH) hydrolysis, and **c**, the MPH trajectory toward parathion-methyl (PTM) hydrolysis. Key steps with no epistasis ($\epsilon = 0$) or with mean epistasis ($\epsilon = \bar{\epsilon}$) are highlighted along each trajectory for comparison. **b**, Structure of WT PTE (PDB: 4PCP) overlaid with an evolved PTE variant containing a 2-naphthyl hexanoate (2NH) transition state analogue (PDB: 43ET). **d**, Structure of a derived MPH variant (PDB: 1P9E) with molecular docking of parathion-methyl (PTM)²⁴. Red bars represent positive epistasis, while blue bars indicate negative epistasis. Idiosyncratic (ϵ) and mean epistasis ($\bar{\epsilon}$) are displayed for each interaction. Blue and red residues represent destabilizing and stabilizing mutations with respect to idiosyncratic epistasis, respectively. The blue and red color coding is consistent across all panels.

Discussion

In this study, we demonstrated high levels of idiosyncrasy in both single mutational effects and epistasis across 45 combinatorial landscapes representing seven enzymes. The success in predicting the functional effects within combinatorial landscapes by the global model, using epistatic coefficients as low as the 3rd order (Fig. 4c), may suggest that fitness landscapes (up to seven mutational combinations) can be explained by relatively simple interaction terms^{16,17}. In this study, however, we find that the global view is not always useful to deconvolute specific evolutionary trajectories, because each mutational path can contain idiosyncrasy that permits or restricts the accessibility of certain genotypes. The idiosyncrasy is inherently a by-product of higher-order epistasis; thus, higher-order epistasis should not just be perceived as an additional epistatic contribution stemming from three or more characterised mutations, but more

importantly, as the source of evolutionary idiosyncrasy. Indeed, we observed high idiosyncrasy even in the 3rd and 4th orders of epistasis suggesting that epistasis at the 5th – and higher – order(s) is creating strong idiosyncrasy which, in turn, leads to a strong deviation between the epistasis arising in a particular genotype and its global (or average) effect (Fig. 3c and Extended Data Figs. 3-4). Thus, as a protein gradually accumulates mutations, the single mutational effects and epistasis are expected to be steadily idiosyncratic, and therefore predicting their evolutionary behaviours at any given genotype is challenging.

The underlying biophysical basis of idiosyncrasy has yet to be uncovered. We suspect that the observed epistatic interactions are partitioned into subnetworks within the grand interaction network of the protein, including the protein's ligands, cofactors, and substrates^{1,36–38}. If the modularity of protein epistatic networks is indeed a key factor in dictating idiosyncrasy, and therefore a major component of understanding higher-order epistasis, it is essential to combine epistatic analysis with techniques probing the molecular and structural bases of these networks³⁹, *e.g.*, using tools such as molecular dynamics (MD)^{38,40} and nucleic magnetic resonance (NMR)⁴¹, particularly in the context of interactions through other molecules¹. By deepening our understanding of these parameters, we may be able to account for idiosyncrasy *a priori* and smoothen the apparent noise in our protein fitness landscapes.

Our observations also provide important implications for the engineering of new proteins. Recent advances in high-throughput mutational characterisation approaches, such as deep mutational scanning (DMS), provide access to a large mutational dataset around the WT background but are generally limited to single and double mutants^{42,43}. However, engineered targets generally stray further from their genotype, and, as we have shown, local information is insufficient for the accurate prediction of evolutionary intermediates and endpoints (Fig. 4c and 5a). In some instances, this caveat has been overcome using *in silico* techniques employing co-evolutionary based models such as direct coupling analysis⁴⁴, or natural language processing⁴⁵, which demonstrate great success in predicting sequences with enhanced functions. Albeit, such techniques are currently unsatisfactory when it comes to predictions of evolutionary paths towards non-native ligands and substrates. We acknowledge that there are more sophisticated machine-learning (ML) based models that may provide greater success in functional prediction than the ones used in this study^{46,47}. However, explicit incorporation of higher-order epistasis into these models has shown impaired prediction, possibly due to the incorporation of 'noise' stemming from idiosyncratic effects⁴⁶. Nevertheless, our results advocate attempts to incorporate idiosyncrasy into ML models, or at least warrant an acknowledgement of the ability to which idiosyncrasy can distort predictions. Likewise, the prevalence and impact of idiosyncrasy shown in our study reaffirm the importance of broader fitness landscape exploration, where ML can also be employed⁴⁷. We hope that in the future it may be feasible to use the predictive power of ML models and broader characterisation of fitness landscapes, combined with other biophysical and biochemical inputs such as DMS, structural prediction tools, and biophysical analyses, to gain a deeper understanding of idiosyncrasy and higher-order epistasis, greatly elevating the level of prediction for protein engineering efforts.

Methods

PTE Combinatorial landscape. During the directed evolution of PTE for higher arylesterase activity, we previously identified a cluster of six function-altering mutations³⁰⁻³². Here, we construct one of the combinatorial landscapes analysed in this study. We explored these six positions on the genetic background of WT PTE (64 variants), and tested all of their combinations for activity against 2-naphthyl hexanoate (2NH), as described previously¹³. Briefly, the 64 variants were constructed by site-directed mutagenesis and subcloned into a pET-27-STREP vector¹³. The variants were then transformed into *E. coli* BL21(DE3) carrying the pGro7 plasmid (Takara, Shiga, Japan) for GroEL/ES chaperones co-expression⁴⁸. Variants were individually inoculated in 96-deep well plates containing lysogeny broth (LB) media, 100 µg/mL ampicillin, and 34 µg/mL chloramphenicol, then grown overnight at 30°C. Overnight cultures were transferred to a new deep well plate containing LB, supplemented with 100 µg/mL ampicillin, 34 µg/mL chloramphenicol, 200 µM ZnCl₂, and 0.2% (w/v) arabinose for chaperone co-expression, then induced with 1 mM IPTG. Pellets were lysed with lysis buffer (50 mM Tris-HCl buffer, 100 mM NaCl, pH 7.5, 0.1% (w/v) Triton-X100, 200 µM ZnCl₂, 100 µg/mL lysozyme and 1 µL benzonase (25 U/µL) per 100 mL of lysis buffer). Lysates were incubated with 200 µM 2NH + 1 mM Fast Red and hydrolysis was monitored at 500 nm.

Data processing. Data from various studies were compiled and standardized to a spreadsheet format described previously¹³. The values of reported functions were divided by the WT-background function, or, in the presence of replicates, by the mean of the WT-background functions, then log₁₀ transformed. However, we raised 10 to the power of all the TEM-1 growth rate values from Mira *et al.* before using them in our standard pipeline, as the growth rates from this study are assumed to be additive and not multiplicative. All processed files are provided (Supplementary File 1). We also chose not to apply the non-linear power transform^{19,49}, which we address in the supplementary material (Supplementary File 3).

ΔF and ε calculation. Genotypes were represented by a string of amino acids that underwent mutation, then encoded using '0' for ancestral states, and '1' for derived states at the given amino acid positions. The single mutational effect (ΔF) of a mutation at position *i* was calculated for each genetic background by computing the difference between $F_{i=0}$ and $F_{i=1}$. The mutational transition for the ΔF is denoted with 'x' which represents a transition from '0' to '1', e.g., the ΔF_{x001} represents ΔF of mutating position 1 in the position 4 mutant background and is equal to $F_{1001} - F_{0001}$. Epistasis (ε) between two positions *i* and *j* were calculated for each genetic background by computing the difference between the ΔF_{i=x, j=0} and ΔF_{i=x, j=1}. For example, the ε_{xx00} represents the additional functional contribution of the combination of the first and second position in the WT background and is equal to ΔF_{x100} - ΔF_{x000}. This is equivalent to ΔF_{1x00} - ΔF_{0x00}. As with pairwise interactions, higher-order ε was calculated by taking the difference between ε of the previous order, e.g., ε_{xxx0} = ε_{xx10} - ε_{xx00}.

WT-background model. The function of a genotype was predicted using the WT-background model as a sum of all mutational (ΔF) and epistatic (ε) effects in the WT-background:

$$F = \sum_{i=1}^n \Delta F_i x_i + \sum_{i < j}^n \varepsilon_{ij} x_i x_j + \dots \quad (2)$$

Where *i* (and *j*) represents the index of residue position, x_i is either '0' or '1' depending on the mutational state of the residue in the given genotype, and *F* is the function. This is analogous to the biochemical view

of epistasis from Poelwijk *et al.*¹⁶, however, our model uses F values that do not represent ΔG of the protein or enzymatic reaction.

Global model. The linear regression model, or global model, employed for epistatic analysis was the same as those used previously^{13,14,25}. Briefly, mutations at n residue positions were annotated with variables ‘-1’ for the ancestral state or ‘1’ for the derived state. These were used as x variables in the linear model, F is the \log_{10} transformed and WT-normalized function of the variant. The linear model was constructed such that:

$$F = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i < j}^n \beta_{ij} x_i x_j + \dots + \epsilon \quad (3)$$

Where i (and j) represents the index of residue position, x_i is either ‘-1’ or ‘1’ depending on the mutational state of the residue in the given genotype, β are the linear coefficients, and ϵ is the error¹⁶.

Absolute Error (AE) calculation for the endpoint variant. For each of the 45 combinatorial landscapes, the endpoint was selected as the genotype furthest mutationally removed from the WT-background. AE of the prediction from each order of the WT background and linear model was calculated using the following formula:

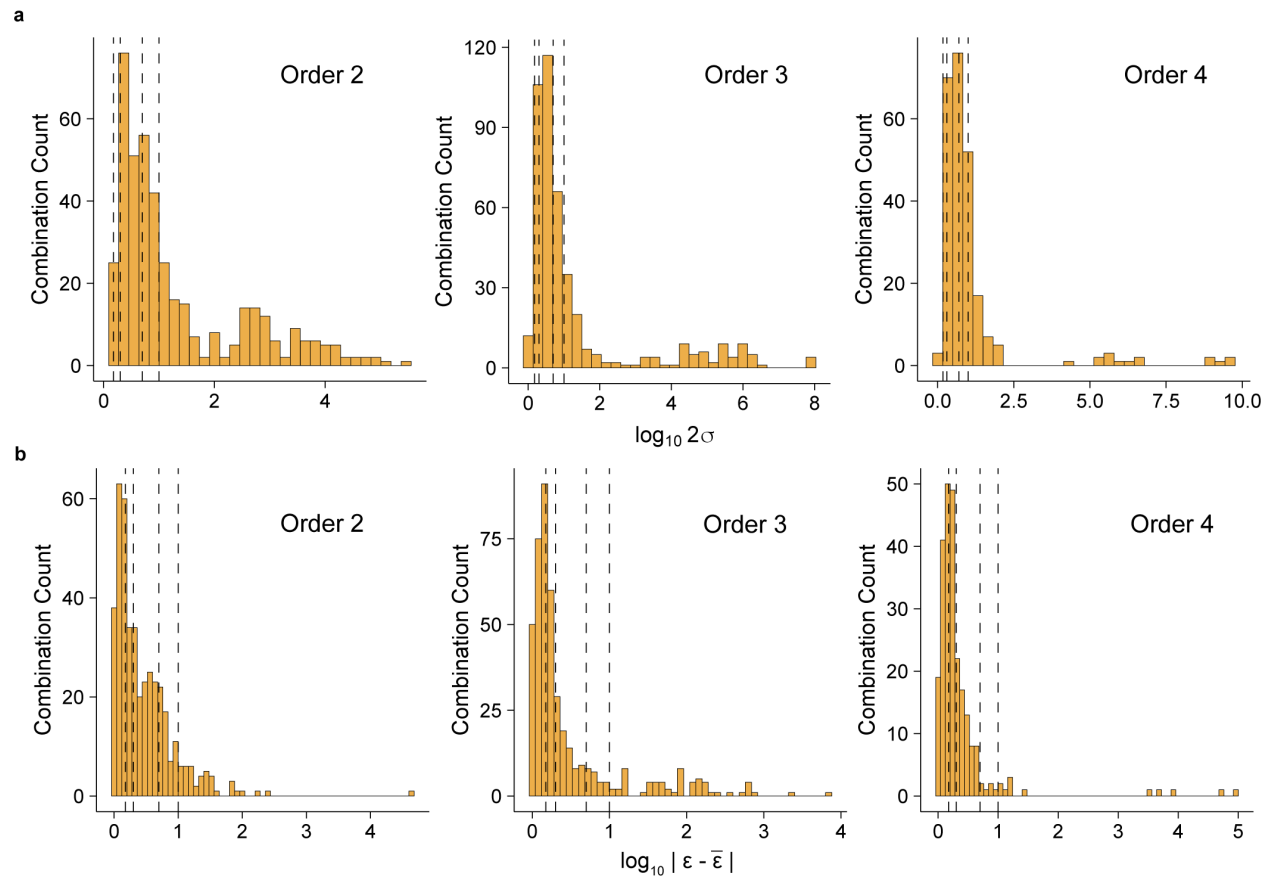
$$AE = |F_{obs} - F_{pred}| \quad (4)$$

Where F_{obs} is the \log_{10} WT-normalized function of the final variant in the landscape, and F_{pred} is the \log_{10} WT-normalized predicted function of the final variant. The mean and median of the AEs in Fig. 4c represents the arithmetic mean and median of the AEs for each model and order pairing.

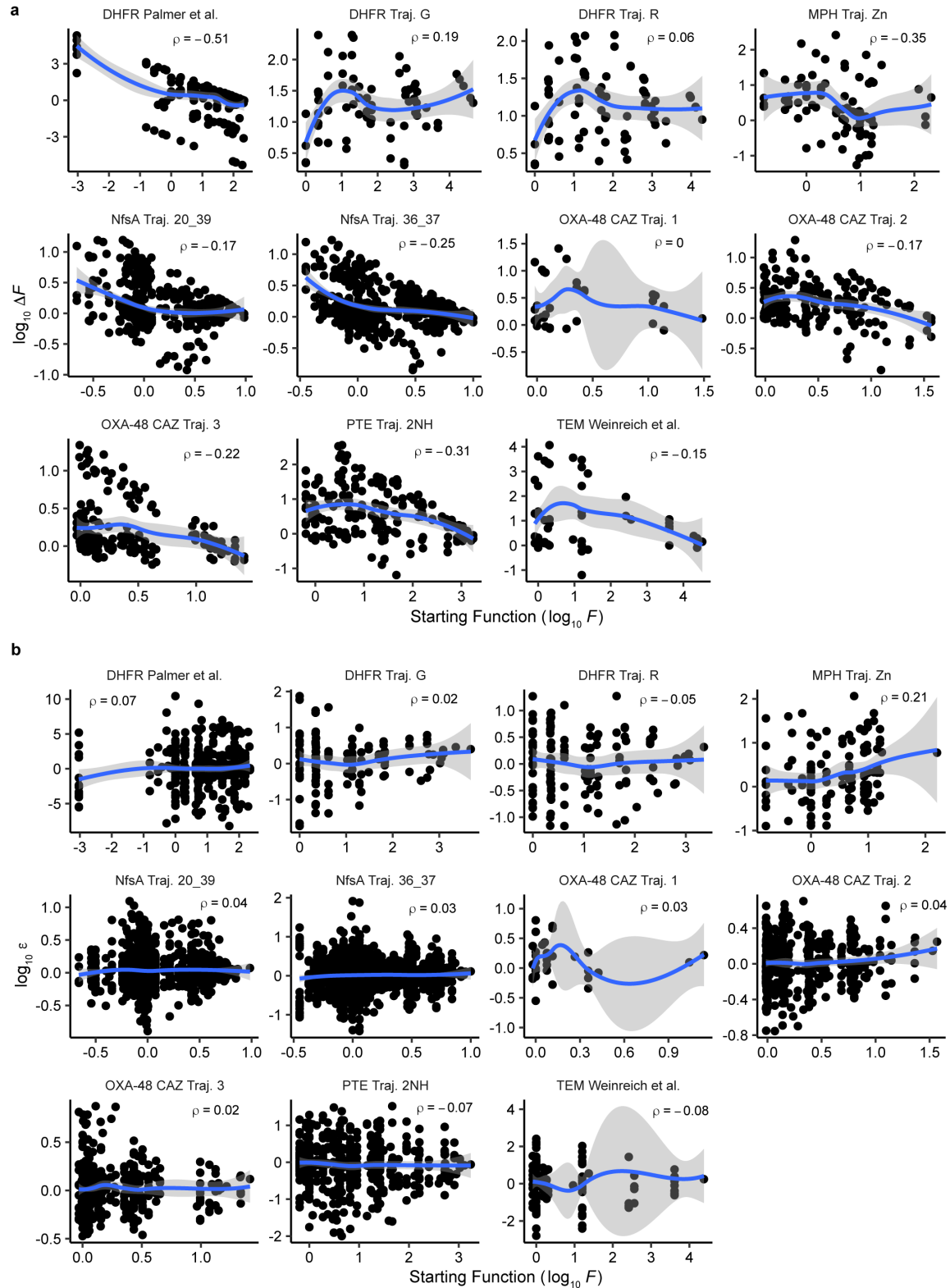
Code and Data Availability

Data for all combinatorial landscapes are provided in \log_{10} -transformed WT-normalized format (Supplementary Data 1). Processed data for functional contributions (Supplementary Data 2) and epistasis (Supplementary Data 3) are also available. Scripts for individual combinatorial landscape analysis and the global statistical analysis are publicly available via GitHub at https://github.com/karolbuda/epistasis_analysis_v2. The scripts for plotting epistasis on PDB structures were adapted from Miton *et al.*¹³. The scripts utilize the R language (<https://www.R-project.org/>), along with R packages: tidyverse (<https://CRAN.R-project.org/package=tidyverse>), igraph (<https://CRAN.R-project.org/package=igraph>), gtools (<https://CRAN.R-project.org/package=gtools>), e1071 (<https://CRAN.R-project.org/package=e1071>), svglite (<https://CRAN.R-project.org/package=svglite>), ggrepel (<https://CRAN.R-project.org/package=ggrepel>), ggpubr (<https://CRAN.R-project.org/package=ggpubr>), and knitr (<https://yihui.org/knitr>).

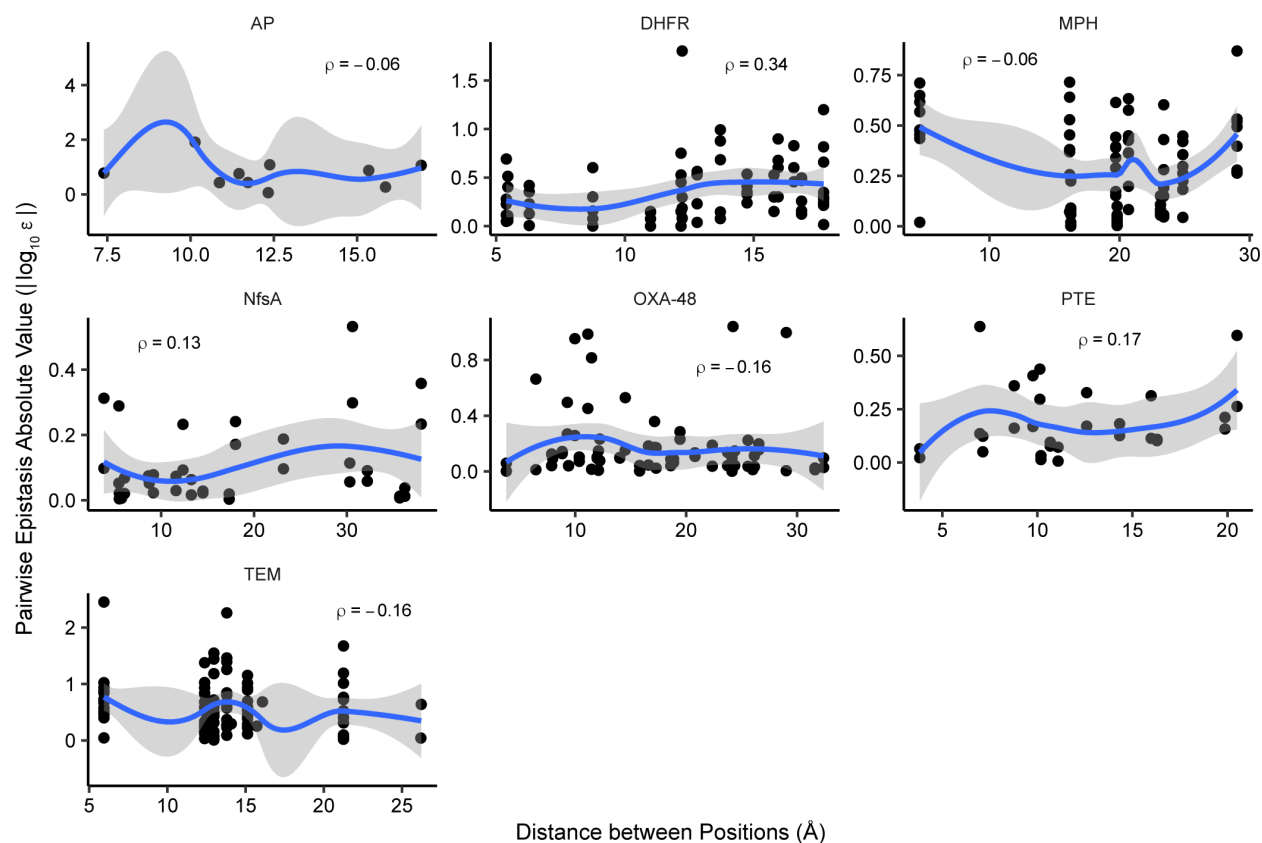
Extended Data



Extended Data Fig. 1 | Expanded histograms from Figs. 3a and 3c. a, Distribution of $\log_{10} 2\sigma$ of ε at each order for all mutational combinations, with annotated dashed lines representing 1.5-fold, 2-fold, 5-fold, and 10-fold significance thresholds, respectively. **b**, Distribution of the WT-background idiosyncratic epistasis quantified via the absolute difference of ε from $\bar{\varepsilon}$ at each combination. Dashed lines represent 1.5-, 2-, 5-, and 10-fold significance thresholds, respectively.



Extended Data Fig. 2 | Trends in mutational and epistatic effects with respect to a genotype's starting function. a, Relationship between starting function and mutational functional contribution (ΔF) for 11 adaptive landscapes. **b,** Relationship between starting function and epistasis (ϵ) for 11 adaptive landscapes. Lines represent local regressions using the LOESS method. Rho (ρ) represents the Spearman correlation coefficient for each landscape.



Extended Data Fig. 3 | Trends in pairwise epistasis with respect to alpha carbon distance between the two positions. Only epistatic values from the 11 adaptive landscapes were used. Structures were obtained from PDB IDs (3TG0, 3CXK, 6A2K, 5HIF, 7NMP, 3HBR, 4PCP, 1XPB) for AP, *E.coli* DHFR, *P. falciparum* DHFR, MPH, NfsA, OXA-48, PTE, and TEM respectively. Lines represent local regressions using the LOESS method. Rho (ρ) represent the spearman correlation coefficient for each landscape.

Extended Data Table 1 | Sign variation of the positions' (1st order) and combinations' (2nd, 3rd, and 4th orders) ΔF 's and ϵ

Order	Homogeneous			Semi-Homogeneous		Heterogeneous
	Negative	Neutral	Positive	Negative-Neutral	Neutral-Positive	Negative-Positive
1	1.4% (3/214)	0.9% (2/214)	9.8% (21/214)	7.5% (16/214)	16.4% (35/214)	64.0% (137/214)
2	1.2% (5/419)	1.2% (5/419)	3.1% (13/419)	8.8% (37/419)	15.3% (64/419)	70.4% (295/419)
3	3.7% (16/438)	2.3% (10/438)	11.9% (52/438)	14.6% (64/438)	13.5% (59/438)	54.1% (237/438)
4	4.5% (11/245)	2.0% (5/245)	4.1% (10/245)	20.4% (50/245)	17.1% (42/245)	51.8% (127/245)

Extended Data Table 2 | Percent of positions (1st order) and combinations (2nd, 3rd, and 4th order) that show a 2σ for ΔF or ϵ above the listed thresholds.

Order	Threshold			
	1.5-fold	2-fold	5-fold	10-fold
1	98.6% (211/214)	95.8% (205/214)	70.1% (150/214)	53.3% (114/214)
2	98.6% (413/419)	92.4% (387/419)	60.1% (252/419)	40.6% (170/419)
3	95.9% (420/438)	85.4% (374/438)	45.4% (199/438)	29.7% (130/438)
4	98.0% (240/245*)	83.3% (204/245*)	49.8% (122/245*)	24.9% (61/245*)

* 4th order considers 245 out of the total 267 combinations as some landscapes have a single 4th order data point and a 2σ cannot be computed

Extended Data Table 3 | Percent of positions and combinations that show a WT ΔF or WT ϵ to mean ΔF or mean ϵ deviation above the listed thresholds.

Order	Threshold			
	1.5-fold	2-fold	5-fold	10-fold
1	67.8% (145/214)	51.4% (110/214)	19.6% (42/214)	6.5% (14/214)
2	64.4% (270/419)	51.6% (216/419)	21.7% (91/419)	6.4% (27/419)
3	55.7% (244/438)	34.5% (151/438)	18.9% (83/438)	7.3% (32/438)
4	62.4% (153/245)	33.9% (83/245)	6.9% (17/245)	4.9% (12/245)

Extended Data Table 4 | Percent of genotypes that show a ΔF or ϵ to mean ΔF or mean ϵ deviation above the listed thresholds.

Order	Threshold			
	1.5-fold	2-fold	5-fold	10-fold
1	59.0% (2396/4064)	40.5% (1646/4064)	14.8% (601/4064)	8.1% (331/4064)
2	52.8% (2466/4672)	34.9% (1630/4672)	13.4% (624/4672)	7.7% (358/4672)
3	49.3% (1467/2976)	30.7% (913/2976)	11.6% (345/2976)	7.6% (227/2976)
4	60.0% (627/1120)	34.6% (387/1120)	9.0% (101/1120)	5.0% (56/1120)

Extended Data Table 5 | Mean and median absolute error of the models' prediction for the derived variant at each model order.

Model	1 st Order	2 nd Order	3 rd Order	4 th Order
Linear	2.9-fold 2.3-fold	2.0-fold 1.8-fold	1.3-fold 1.2-fold	1.1-fold 1.1-fold
WT-background	20.5-fold 8.0-fold	106.2-fold 63.7-fold	51.1-fold 9.0-fold	25.5-fold 4.5-fold

Extended Data Table 6 | WT-background model absolute error in predicted log-transformed function ($\log_{10} F$) of each variant in the most accessible path by model order.

Predicted Variant Mutations	Trajectory	Order					
		1 st	2 nd	3 rd	4 th	5 th	6 th
2	DHFR IC ₇₅	0.03	-	-	-	-	-
3		0.46	0.12	-	-	-	-
2	DHFR G-Traj.	0.61	-	-	-	-	-
3		1.26	0.02	-	-	-	-
4		2.14	1.90	0.36	-	-	-
5		3.17	3.04	2.02	0.40	-	-
2	DHFR R-Traj	0.61	-	-	-	-	-
3		1.26	0.02	-	-	-	-
4		1.84	0.83	0.17	-	-	-
5		2.43	2.63	1.17	0.50	-	-
2	MPH Zn PTM	0.06	-	-	-	-	-
3		0.69	0.00	-	-	-	-
4		0.55	1.90	0.23	-	-	-
5		0.88	3.38	0.85	0.70	-	-
2	NfsA 20_39	0.14	-	-	-	-	-
3		0.32	0.01	-	-	-	-
4		0.44	0.09	0.07	-	-	-
5		0.60	1.17	0.27	0.23	-	-
2	NfsA 36_37	0.13	-	-	-	-	-
3		0.32	0.17	-	-	-	-
4		0.47	0.46	0.54	-	-	-
5		0.81	0.12	0.88	0.91	-	-
6		0.87	0.33	1.53	1.42	0.74	-
7		0.91	0.2	0.71	2.23	3.50	1.91
2		OXA-48 CAZ Traj. 1	0.80	-	-	-	-
3	1.07		0.55	-	-	-	-
4	1.08		0.32	0.27	-	-	-
2	OXA-48 CAZ Traj. 2	0.03	-	-	-	-	-
3		0.18	0.34	-	-	-	-
4		0.63	0.62	0.38	-	-	-
5		0.68	0.70	0.13	0.13	-	-
6		0.61	0.45	0.80	0.72	0.24	-
2	OXA-48 CAZ Traj. 3	0.50	-	-	-	-	-
3		0.84	0.21	-	-	-	-
4		0.88	0.50	0.19	-	-	-
2	PTE 2NH	0.03	-	-	-	-	-
3		0.17	1.00	-	-	-	-
4		0.58	0.95	0.88	-	-	-
2	TEM MIC	2.23	-	-	-	-	-
3		3.00	1.47	-	-	-	-
4		3.23	4.16	1.93	-	-	-
5		3.38	4.24	1.27	0.27	-	-

Acknowledgements

We thank D. Anderson and members of the Tokuriki lab for valuable comments. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN 2017-04909) and Human Frontier Science Program (HFSP) Research Grant (RGP0054/2020).

Author Contributions

K.B. and N.T. conceived of and designed the study. K.B. and N.T. performed all statistical analyses. C.M.M. performed activity assays for the PTE combinatorial landscape. All authors contributed to writing the paper.

Competing Interests

The authors declare no competing interests.

References

1. Miton, C. M., Buda, K. & Tokuriki, N. Epistasis and intramolecular networks in protein evolution. *Curr Opin Struc Biol* 69, 160–168 (2021).
2. Kaltenbach, M. & Tokuriki, N. Dynamics and constraints of enzyme evolution. *J Exp Zoology Part B Mol Dev Evol* 322, 468–487 (2014).
3. Lehner, B. Molecular mechanisms of epistasis within and between genes. *Trends Genet* 27, 323–331 (2011).
4. Visser, J. A. G. M. de, Elena, S. F., Fragata, I. & Matuszewski, S. The utility of fitness landscapes and big data for predicting evolution. *Heredity* 121, 401–405 (2018).
5. Kondrashov, D. A. & Kondrashov, F. A. Topological features of rugged fitness landscapes in sequence space. *Trends Genet* 31, 24–33 (2015).
6. Hartman, E. C. & Tullman-Ercek, D. Learning from protein fitness landscapes: a review of mutability, epistasis, and evolution. *Curr Opin Syst Biology* 14, 25–31 (2019).
7. Weinreich, D. M., Watson, R. A. & Chao, L. PERSPECTIVE: SIGN EPISTASIS AND GENETIC CONSTRAINT ON EVOLUTIONARY TRAJECTORIES. *Evolution* 59, 1165–1174 (2005).
8. Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M. & Tans, S. J. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* 445, 383–386 (2007).
9. Visser, J. A. G. M. de & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nat Rev Genet* 15, 480–490 (2014).
10. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Bio* 10, 866–876 (2009).
11. Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins. *Science* 312, 111–114 (2006).
12. Anderson, D. W., McKeown, A. N. & Thornton, J. W. Intermolecular epistasis shaped the function and evolution of an ancient transcription factor and its DNA binding sites. *Elife* 4, e07864 (2015).
13. Miton, C. M., Chen, J. Z., Ost, K., Anderson, D. W. & Tokuriki, N. Statistical analysis of mutational epistasis to reveal intramolecular interaction networks in proteins. *Methods Enzymol* 643, 243–280 (2020).
14. Yang, G. *et al.* Higher-order epistasis shapes the fitness landscape of a xenobiotic-degrading enzyme. *Nat Chem Biol* 15, 1120–1128 (2019).
15. Weinreich, D. M., Lan, Y., Jaffe, J. & Heckendorn, R. B. The Influence of Higher-Order Epistasis on Biological Fitness Landscape Topography. *J Stat Phys* 172, 208–225 (2018).
16. Poelwijk, F. J., Krishna, V. & Ranganathan, R. The Context-Dependence of Mutations: A Linkage of Formalisms. *Plos Comput Biol* 12, e1004771 (2016).
17. Sailer, Z. R. & Harms, M. J. High-order epistasis shapes evolutionary trajectories. *Plos Comput Biol* 13, e1005541 (2017).
18. Reddy, G. & Desai, M. M. Global epistasis emerges from a generic model of a complex trait. *Elife* 10, e64740 (2021).
19. Otwinowski, J., McCandlish, D. M. & Plotkin, J. B. Inferring the shape of global epistasis. *Proc National Acad Sci* 115, 201804015 (2018).
20. Lyons, D. M., Zou, Z., Xu, H. & Zhang, J. Idiosyncratic epistasis creates universals in mutational effects and evolutionary trajectories. *Nat Ecol Evol* 1–9 (2020) doi:10.1038/s41559-020-01286-y.
21. Bakerlee, C. W., Ba, A. N. N., Shulgina, Y., Echenique, J. I. R. & Desai, M. M. Idiosyncratic epistasis leads to global fitness-correlated trends. *Science* 376, 630–635 (2022).
22. Fröhlich, C. On the Evolvability of OXA-48. (UiT The Arctic University of Norway, 2021).
23. Palmer, A. C. *et al.* Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. *Nat Commun* 6, 7385 (2015).
24. Hall, K. R. *et al.* Intracellular complexities of acquiring a new enzymatic function revealed by mass-randomisation of active site residues. *Elife* 9, e59081 (2020).
25. Anderson, D. W., Baier, F., Yang, G. & Tokuriki, N. The adaptive landscape of a metallo-enzyme is shaped by environment-dependent epistasis. *Nat Commun* 12, 3867 (2021).
26. Lozovsky, E. R., Daniels, R. F., Heffernan, G. D., Jacobus, D. P. & Hartl, D. L. Relevance of Higher-Order Epistasis in Drug Resistance. *Mol Biol Evol* 38, msaa196- (2020).
27. Tamer, Y. T. *et al.* High-order epistasis in catalytic power of dihydrofolate reductase gives rise to a rugged fitness landscape in the presence of trimethoprim selection. *Mol Biol Evol* 36, 1533–1550 (2019).
28. Mira, P. M. *et al.* Rational Design of Antibiotic Treatment Plans: A Treatment Strategy for Managing Evolution and Reversing Resistance. *Plos One* 10, e0122283 (2015).
29. Sunden, F., Peck, A., Salzman, J., Ressler, S. & Herschlag, D. Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. *Elife* 4, e06181 (2015).
30. Tokuriki, N. *et al.* Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat Commun* 3, 1257 (2012).

31. Campbell, E. *et al.* The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* 12, 944–950 (2016).
32. Wyganowski, K. T., Kaltenbach, M. & Tokuriki, N. GroEL/ES Buffering and Compensatory Mutations Promote Protein Evolution by Stabilizing Folding Intermediates. *J Mol Biol* 425, 3403–3414 (2013).
33. Card, K. J., Jordan, J. A. & Lenski, R. E. Idiosyncratic variation in the fitness costs of tetracycline-resistance mutations in *Escherichia coli*. *Evolution* 75, 1230–1238 (2021).
34. Schenk, M. F., Szendro, I. G., Salverda, M. L. M., Krug, J. & Visser, J. A. G. M. de. Patterns of Epistasis between Beneficial Mutations in an Antibiotic Resistance Gene. *Mol Biol Evol* 30, 1779–1787 (2013).
35. Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci* 25, 1204–1218 (2016).
36. Yan, W. *et al.* The construction of an amino acid network for understanding protein structure and function. *Amino Acids* 46, 1419–1439 (2014).
37. Xiangmao, M., Wenkai, L., Xiaoqing, P., Yaohang, L. & Min, L. Protein interaction networks: centrality, modularity, dynamics, and applications. *Frontiers of Computer Science* doi:10.1007/s11704 - 020 - 8179 - 0.
38. Bunzel, H. A. *et al.* Evolution of dynamical networks enhances catalysis in a designer enzyme. *Nat Chem* 13, 1017–1022 (2021).
39. Morrison, A. J., Wonderlick, D. R. & Harms, M. J. Ensemble epistasis: thermodynamic origins of nonadditivity between mutations. *Genetics* 219, iyab105 (2021).
40. Maria-Solano, M. A., Serrano-Hervás, E., Romero-Rivera, A., Iglesias-Fernández, J. & Osuna, S. Role of conformational dynamics in the evolution of novel enzyme function. *Chem Commun* 54, 6622–6634 (2018).
41. Kurzbach, D. Network representation of protein interactions: Theory of graph description and analysis. *Protein Sci* 25, 1617–1627 (2016).
42. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat Methods* 11, 801–807 (2014).
43. Fernandez-de-Cossio-Diaz, J., Uguzzoni, G. & Pagnani, A. Unsupervised inference of protein fitness landscape from deep mutational scan. *Mol Biol Evol* 38, msaa204 (2020).
44. Russ, W. P. *et al.* An evolution-based model for designing chorismate mutase enzymes. *Sci New York N Y* 369, 440–445 (2020).
45. Shin, J.-E. *et al.* Protein design and variant prediction using autoregressive generative models. *Nat Commun* 12, 2403 (2021).
46. Sailer, Z. R. *et al.* Inferring a complete genotype-phenotype map from a small number of measured phenotypes. *Plos Comput Biol* 16, e1008243 (2020).
47. Wu, Z., Kan, S. B. J., Lewis, R. D., Wittmann, B. J. & Arnold, F. H. Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc National Acad Sci* 116, 201901979 (2019).
48. Tokuriki, N. & Tawfik, D. S. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* 459, 668–673 (2009).
49. Sailer, Z. R. & Harms, M. J. Detecting High-Order Epistasis in Nonlinear Genotype-Phenotype Maps. *Genetics* 205, 1079–1088 (2017).