

Enzyme engineering: Reaching the maximal catalytic efficiency peak

Moshe Goldsmith and Dan S. Tawfik

Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

Abstract

The practical need for highly efficient enzymes presents new challenges in enzyme engineering, in particular, the need to improve catalytic turnover (k_{cat}) or efficiency (k_{cat}/K_M) by several orders of magnitude. However, optimizing catalysis demands navigation through complex and rugged fitness landscapes, with optimization trajectories often leading to strong diminishing returns and dead-ends. When no further improvements are observed in library screens or selections, it remains unclear whether the maximal catalytic efficiency of the enzyme (the catalytic ‘fitness peak’) has been reached; or perhaps, an alternative combination of mutations exists that could yield additional improvements. Here, we discuss fundamental aspects of the process of catalytic optimization, and offer practical solutions with respect to overcoming optimization plateaus.

Highlights

- Optimization plateaus are common when engineering enzymes for higher catalytic efficiency.
- These plateaus relate to fundamental properties of evolutionary fitness landscapes.
- Marginal protein stability is a common cause of plateauing that can be easily overcome.
- Activity tradeoffs and epistatic effects are other causes of optimization plateaus.

Introduction

It is commonly assumed that all enzymes were born poor catalysts and were subsequently optimized by evolution. Laboratory experiments that mimic this process have now become a matter of routine. In some cases, these experiments attempt to reproduce the evolutionary emergence of a natural enzyme from its putative ancestor. However, in most cases, the target of enzyme optimization is primarily applicative - to create a highly active and stable enzyme that can catalyze the target reaction in a non-biological setting. Natural enzymes present two common limitations: First, with few exceptions, natural enzymes exhibit low catalytic efficiency with non-cognate substrates that are typically application-relevant (*i.e.*, substrates that differ from the enzyme's natural substrate yet are promiscuously transformed by it). Second, natural enzymes exhibit low protein stability especially under applicative conditions. The feasibility and cost of application depend on the catalytic turnover of the enzymes being high. To this end, directed evolution, and other protein engineering methods such as computational design, are used to optimize enzymes for a variety of practical applications spanning from organic synthesis to therapeutics (for recent reviews, see Refs. [1-5]).

Here, we summarize several key aspects regarding the laboratory optimization of enzymatic traits, and specifically of catalytic efficiency (k_{cat}/K_M , or k_{cat} for enzymes working under substrate saturation [6]). The last decade has seen a leap in the understanding of how enzymes evolve. However, several key questions still prevail, especially with respect to how enzymes can be optimized toward high, let alone maximal catalytic efficiency (in fact, what 'high' or 'maximal' means is a complex issue, as discussed below). We focus on optimization of catalytic efficiency, primarily by directed evolution. There are fundamental differences between evolution in nature and in the laboratory. Nonetheless, lessons from natural evolution can be implemented in enzyme engineering, and the latter also teaches us about the former.

The optimization challenge

The catalytic efficiencies of natural enzymes with non-cognate substrates can be extremely low; reported k_{cat}/K_M values as low as $1 \text{ M}^{-1}\text{s}^{-1}$ are not an exception (e.g. [7-10]). In contrast, the average catalytic efficiency value (k_{cat}/K_M) of natural enzymes with their cognate substrates is $\sim 10^5 \text{ M}^{-1}\text{s}^{-1}$, and some enzymes approach $10^9 \text{ M}^{-1}\text{s}^{-1}$ [11]. Thus, the gap between the catalytic efficiency with a cognate substrate versus a promiscuous, non-cognate substrate may span several orders of magnitude. Natural

evolution can readily bridge such gaps. For example, xenobiotics such as the pesticide metabolite paraoxon are promiscuously hydrolyzed by many natural enzymes, although at very different rates. Bacterial lactonases known as PLLs exhibit promiscuous paraoxonase activities with k_{cat}/K_M values that span over 4 orders-of-magnitude (from 0.5 up to $4 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$) [12]. A natural paraoxonase named PTE (phosphotriesterase), had diverged from an unidentified PLL. This enzyme evolved in soil bacteria, in a matter of few decades, to hydrolyze paraoxon with a rate that approaches diffusion limit ($k_{\text{cat}}/K_M \sim 10^8 \text{ M}^{-1}\text{s}^{-1}$) [13].

Directed evolution and computational redesign of natural enzymes have proven capable of bridging equally large gaps; possibly even greater ones when applying starting points that exhibit no detectable activity with the target substrate (e.g. Refs. [14,15]). However, most publications describe only modest improvements of up to two orders of magnitude in k_{cat}/K_M (**Figure 1**). Large improvements are rare, and increases in k_{cat}/K_M of $\geq 10^4$ -fold comprise only $\sim 5\%$ of our literature sampling (**Figure 1**). Improving an enzyme's catalytic efficiency with a non-cognate substrate by an order of magnitude or two typically requires only a few rounds of directed evolution, especially when the initial catalytic efficiency is relatively low. In contrast, efforts to bridge large gaps by directed evolution are likely to encounter diminishing returns and optimization plateaus, and thus require many rounds of directed evolution with no guarantee of success. The correlation between fold-improvement and the number of introduced mutations is not strictly linear, but improvements of above 1,000-fold typically demand at least 10 mutations (**Figure 2**).

The landscape of evolutionary optimizations

The search for a combination of beneficial mutations that would yield an optimized enzyme is usually depicted as a sequential advancement in protein sequence-space, where each step is associated with a change in fitness. An in-depth discussion of this topic is beyond the scope of this review, but for the purposes of this discussion, fitness landscapes can be illustrated by a simplified 3D space (**Box 1**). In the simplest scenario, there exists a trajectory that is both continuous and gradual – an uphill climb to the 'fitness peak' that comprises a series of mutations, each of which provides a distinct fitness advantage (**Box 1, trajectory A**).

The fitness peak represents the maximal possible catalytic efficiency of an enzyme for a particular reaction and substrate in a given region of sequence space. In theory, an unlimited exploration of protein sequence space will reveal the 'global fitness

peak', which represents the maximal k_{cat}/K_M value possible for a given enzyme, reaction and substrate. However, the number and heights of local fitness peaks that exist for a particular enzyme and substrate are unknown, and at present are also impossible to predict. Given the vastness of the theoretical sequence space, only a minute portion of it is accessible to natural evolution, let alone to laboratory evolution. The size of sequence-space explored by natural evolution is restricted by factors such as the frequency of mutations, their types and the size of the evolving population. Similarly, in laboratory evolution, the diversity of the applied gene library (i.e. the number of mutations per gene, their types, the number of variants per library), and the throughput of the screening/selection method and its stringency, dictate what fraction of sequence space will be explored per round.

How catalytically efficient can a given enzyme become?

The number of natural enzymes that have reached their maximal possible catalytic efficiency (i.e. the global fitness peak) is unknown. Only so-called "perfect enzymes", whose k_{cat}/K_M values approach the physical limit of diffusion rate ($\geq 10^9 \text{ M}^{-1}\text{s}^{-1}$), may be assumed to have done so. However, for more than 98% of natural enzymes, the k_{cat}/K_M are ≥ 10 times lower than diffusion limit [16] and the average literature k_{cat}/K_M value is 4 orders of magnitude lower ($\sim 10^5 \text{ M}^{-1}\text{s}^{-1}$) [11]. A large number of these far-from-perfect enzymes have probably reached their optimal catalytic efficiency in their natural cellular context (i.e. local fitness peaks). This is because *in vivo*, the catalytic efficiency of enzymes is restricted by various flux-balance factors [17,18] such as avoiding overproduction of toxic products or rapid depletion of substrates used by other enzymes. However, many natural enzymes are likely to be far from maximal performance, certainly *in vitro*. For example, enzymes involved in secondary metabolism typically show lower catalytic efficiency compared to enzymes in core metabolism, probably due to relatively recent emergence and weak selection pressures [11]. The fact that most natural enzymes are far from maximal performance is also evident from their configurational stability (the ability to fold and maintain the native fold). Numerous enzymes have been stabilized in the laboratory while maintaining or even increasing catalytic efficiency [19-21].

The implication of the above is that during enzyme optimization, unless a diffusion-rate limit has been reached, there is no way of telling whether a higher catalytic efficiency could be obtained for the evolving enzyme or not. This point is further discussed in the section titled *Optimization landscapes - local versus global fitness*

peaks.

Diminishing returns

The rugged shape of fitness landscapes limits sequence explorations to specific trajectories. Within a given trajectory, even if it is continuous and gradual (trajectory **A Box 1**), the fitness increase per mutation is rarely constant. Foremost, evolutionary optimizations are subject to diminishing returns – the early mutations give rise to relatively large improvements (typically, 5 to 10-fold increases in k_{cat}/K_M), but as the optimization progresses, the fold-improvement in catalytic efficiency per newly added mutation decreases.

Diminishing returns with respect to catalytic efficiency were systematically measured in a trajectory leading from a bacterial phosphotriesterase (PTE) to an aryl, carboxy-esterase [22]. An initially weak, promiscuous activity with a non-cognate aryl-ester substrate was improved nearly 10^5 -fold. This demanded 18 rounds of mutagenesis and screening, whereby, on average, each mutation improved k_{cat}/K_M by a factor of ~ 2 . However, whereas the first 4 mutations gave an overall improvement of $>10^3$ -fold (an improvement factor of ~ 5 per mutation), the improvement factor for the last mutations was ~ 0.18 (**Figure 3**). The first mutations that induce large improvements occurred within the active site, reshaping it to accommodate the new substrate and fine-tune the catalytic machinery. The late mutations were typically 2nd and 3rd shell. These mutations had a role in compensating the destabilizing effects of the 1st shell mutations and reinforcing their effects, and also in funneling the conformational ensemble toward the catalytically productive conformations [22,23]. While these latter mutations are critical, their contribution to catalytic efficiency per se is small, or sometimes nil [22]. Equally strong diminishing returns were similarly encountered when the same enzyme, PTE, was evolved toward hydrolysis of the nerve agent VX [24]. Overall, a $\sim 5 \times 10^3$ -fold improvement in k_{cat}/K_M for hydrolysis of the toxic isomer of VX was obtained following 13 rounds of directed evolution (combining rationally and computationally targeted libraries and random mutagenesis). The first 500-fold improvement was achieved with 5 rounds and 7 mutations (3 within the active-site, and 4 peripheral mutations), whereas the remaining 10-fold demanded 8 more rounds and 12 more mutations (4 active-site, and 8 peripheral mutations) (**Figure 3**). Most of the final 10-fold improvement was obtained only after the introduction of 7 stabilizing mutations, which were derived using computational design [25], and introduced in a single step [24].

Local versus global fitness peaks

Diminishing returns result in the need for multiple rounds to achieve large improvements (**Figure 2**) and they also lead to optimization plateaus (i.e. no enhancements in efficiency despite the introduction of new mutations). Arrested improvements in catalytic efficiency may occur earlier than observed in screens or selections because improvements in other properties, such as protein folding and stability, may increase the observed activity without increasing k_{cat}/K_M . On the other hand, screens (let alone selections) exhibit considerable noise. Thus, during advanced rounds of directed evolution the magnitude of improvements may be within experimental noise, resulting in a prematurely observed plateau. Overall, plateaus in the optimization of catalytic efficiency are commonly encountered. The snag is that when improvements plateau at least three different scenarios might apply that demand fundamentally different courses of action.

The 1st scenario is that the fitness peak has been reached, *i.e.*, the evolved variants confer maximal performance given the starting enzyme, reaction, substrate, and the applied exploration of sequence space (as in **trajectory A, Box 1**). However, as discussed above, one rarely knows what is the maximal k_{cat}/K_M or k_{cat} value expected for a given enzyme optimization. Nonetheless, if one assumes that the peak has been reached, yet is too low in activity relative to the desired goal, a plausible course of action is to explore a new starting point, *i.e.*, another enzyme(s) exhibiting weak, promiscuous activity with the same reaction/substrate. Enzymes exhibiting similar promiscuity may significantly differ in their potential for improvement [26].

The 2nd scenario is that the evolutionary trajectory has reached a local plateau (**Box 1, trajectory B**). No further improvements are observed in the screened libraries although, in principle, higher catalytic efficiency could be achieved via this trajectory. The most common cause of transient, local optimization plateaus is a loss of stability due to the accumulation of mutations. Accordingly, “pulling out” from such a local plateau demands stabilizing mutations, and often a combination of several [25,27]. At a minimum, a combination of two mutations is needed to “pull out”, whereby one (or more) of these mutations is neutral or even slightly deleterious on its own. The latter, so-called enabling, or permissive mutations, comprise a critical background for the acquisition of the mutation which enhances catalytic activity. Enabling mutations could be enriched by performing a ‘neutral drift’ – *i.e.*, accumulating mutations despite the fact that they do not increase activity [15]. Drift, however, may not be effective if the fitness peak can only be reached by a rare combination of mutations

[28]. Alternatively, the method by which sequence space is searched could be modified [1]. For example, gene libraries are primarily based on amino acid substitutions, although backbone changes (insertions and deletions of amino acids; InDels) play a critical role in natural enzyme evolution [29] and can be incorporated into libraries [30,31]. Similarly, swapping whole secondary structural elements between homologous proteins may allow access to different regions of sequence space from which the fitness peak is more easily accessible [32].

The 3rd option is that the particular trajectory that had been followed has led to an isolated peak whose height is lower than the that of a nearby fitness peak (**Box 1, trajectory C**). Contrary to the transient local plateau scenario depicted in the 2nd scenario (**trajectory B**), this local peak is separated from the fitness peak by a deep ‘valley’ of much lower fitness. If this 3rd scenario applies, the only way of reaching the fitness peak is going back to the starting point and initiating a new trajectory, preferably via an alternative library making method (see 2nd scenario).

Epistasis, founder mutations and mutational leaps

The shape of fitness landscapes and the evolutionary trajectories they enable result mostly from tradeoffs and epistasis. In a nutshell, epistasis means that the effects of mutations are non-additive, or sequence context dependent. Thus, to maintain a “smooth”, continuous path of improvement in fitness, mutations must occur in a particular order, or even, in combinations of multiple mutations. Epistasis dominates protein evolution (for recent reviews, see Refs [33,34]) and thus has a profound effect on optimization trajectories. In essence, epistasis also is the cause of evolutionary dead-ends, be it a transient, local plateau (**2nd scenario** above; **trajectory B**) or a local peak (**3rd scenario**; **trajectory C**). Local peaks usually occur due to sign (or reciprocal) negative epistasis, whereby two or more mutations that are individually beneficial, become deleterious when combined. Such mutations initiate orthogonal trajectories and are termed ‘founder mutations’ (**Box 1**). The molecular origins of this phenomenon have been studied with the antibiotics resistance enzyme TEM-1 β -lactamase [35].

In extreme cases, an optimization landscape may be so rugged that the fitness peak can only be reached by a “leap” in sequence space, *i.e.*, by simultaneous incorporation of an entire set of mutations (**trajectory D, Box 1**). If the required set includes more than a few mutations, such peaks may be practically inaccessible by laboratory evolution that usually banks on the gradual accumulation of single

mutations. In contrast, computational design may access such peaks by simultaneously incorporating a large number of mutations to gain a new function or improve an existing one (for example see Refs [32,36]). In one case, computational design enabled the simultaneous incorporation of 7 stabilizing mutations needed to 'pull out' of an optimization plateau (**Figure 3**) [25]. Computational methods can also be used for an analysis of correlated substitutions by identifying positions that tend to co-evolve in protein phylogenies, as was done to facilitate protein structural predictions [37]. Sets of correlated substitutions can be incorporated to boost activity and/or stability, or to identify substitutions that were uncoupled during directed evolution, leading to a decrease in stability [38,39]. Restoring such pairs may help restore protein stability.

How frequent are isolated fitness peaks that can only be reached by mutational leaps is largely unknown, however, enzymes redesigned for a completely new activity comprise a good example. One test case is the computational redesign of the active-site of a zinc deaminase to generate a phosphotriesterase [10]. To uncover the minimal set of mutations required to endow phosphotriesterase activity, a library was constructed exploring random combinations of the 8 substitutions in the designed variant at the background of the wild-type deaminase. Screening of this library indicated that a minimal set of 4 mutations was required for the phosphotriesterase activity to reach a detectable level [10]. The likelihood of identifying active quadruple mutants from libraries is exceedingly low (certainly in random mutations libraries), and such a peak is therefore considered inaccessible by directed evolution.

Pulling-out of optimization plateaus

The challenge of improving enzymes past optimization plateaus has not been frequently addressed. For one thing, most published works describe the early stages of optimization and/or do not aim to meet a high-performance efficiency threshold for a particular application (**Figure 1**). It is likely that many enzymes have been evolved toward high catalytic efficiency for industrial applications (for recent examples, see Refs. [1,40-43]), but most of these cases remain unpublished. Nonetheless, a number of endeavors to increase catalytic efficiency by several orders of magnitude have been described [10,14,15,22,39,44-50] and these indicate some general guidelines.

Stability boosts are critical in long-range optimizations. Most mutations are destabilizing, and mutations that alter enzymatic functions are particularly prone to

do so [51]. Using a highly stable starting point enzyme, either laboratory stabilized, or from a hyper-thermophilic organism, would typically enable a longer, continuous progress of improvement. Ancestral inference may also provide highly stable, mutation-tolerant starting points for directed evolution [52-55]. Chaperones can be applied to assist the folding of impaired evolving mutants [51]. However, as the optimization trajectory progresses further, the need for additional compensatory mutations is bound to arise. Compensation may be specific and local, *i.e.*, through residues that are in direct contact with the active-site mutations or non-specific, global, *i.e.*, via mutations that are far away from the active-site and increase the enzyme's overall configurational stability [34] (for a recent examples see Refs. [24,27]).

Local stabilizing mutations may not be easy to predict, but they may appear as beneficial mutations in screens owing to an increase in the levels of soluble, folded enzyme. They can also be introduced into libraries by diversification of 2nd-shell positions that are in contact with 1st shell residues in which mutations occurred in previous rounds (neighbor-joining strategy; [56]). Global compensatory mutations can be readily identified by computational methods and/or by bioinformatics (see Refs [25,54,57-59]). Global mutations tend to rigidify the enzyme's scaffold, thus increasing its mutational tolerance. Increased scaffold rigidity does not usually have a negative effect on enzymatic activity, or on the potential to acquire new or enhanced enzymatic functions [60]. However, one obstacle is that most globally stabilizing mutations provide relatively modest contributions to stability. Hence, a combination of several, or even many mutations is often needed to provide sufficient global stabilization effects. Computational design seems like the best tool to predict a large number of stabilizing mutations. Variants carrying >50 potentially stabilizing mutations were readily identified by computational design, thus providing a large boost in stability [25]. We have recently applied this computational tool, PROSS [25], to continue catalytic optimization past a plateau [24] (**Figure 3**). Computational stabilization resulted not only in a large increase in stability but also in a 2-fold increase in k_{cat}/K_M . Crucially, the stabilized variant could be further evolved to reach the desired catalytic efficiency [24]. Overall, it appears that stability-boosts comprise an essential step for overcoming optimization plateaus, though sometimes not a sufficient one.

Avoiding tradeoffs – simultaneously selecting for two traits may result in tradeoffs. This is not always the case – for example, selection for higher rate may also endow

higher configurational stability or higher selectivity. However, selecting for improved activity with two different substrates may restrict enzyme optimization since one activity may come at the expense of the other, especially, at advanced stages of optimization. For example, we attempted to select a broad-specificity neutralizing enzyme that would simultaneously hydrolyze two nerve agents: VX and RVX. The reaction is identical and the structural differences between these substrates are relatively small. Nonetheless, screening for variants that are improved in both substrates rapidly led to a plateau, and analysis of the evolving variants indicated a tradeoff between these substrates [24].

Higher screening/selection throughput and accuracy can also help pull-out of plateaus (or, perhaps, avoid them altogether). Screens utilize parallel assays of isolated library variants (typically with a chromogenic or fluorogenic substrate) [61]. They provide accuracy and a systematic view of the library's fitness landscape, but are limited in throughput [56]. Selections rely on a direct link between enzymatic activity and cell growth. They act on variant libraries in bulk and typically have higher throughputs than screens. They enable the exploration of a larger sequence space in each round, including and most crucially, combinations of several mutations at a time. Methods for continuous directed evolution allow to rapidly select large repertoires, pursue long mutational trajectories [62,63], and thus obtain large improvements in binding and selectivity [64,65]. On the other hand, selections are harder to fine-tune, and also tend to rapidly funnel and limit the diversity of improved variants. Accurate, high-throughput screens using, for example, microfluidics devices (for recent examples see [7,66]) may comprise a better solution. Such screens may not only explore larger regions of sequence spaces, but can also facilitate the identification of a greater number of slightly beneficial and even neutral mutations that can help traverse local plateaus.

Finally, the application of neutral drift (i.e. purifying selection for the enzyme's native activity) prior to selection for the new activity, may facilitate optimization past a plateau. Neutral drift can be performed using different selection regimes and promotes the accumulation of enabling or compensating mutations that are a prerequisite for mutations that modulate enzymatic activity [67-71]. Weak purifying selection can be applied to only purge inactive variants while retaining slightly deleterious mutations that may become beneficial with subsequent ones [72]. Thus, neutral drift may enable the acquisition of mutations that open novel evolutionary pathways in sequence space.

To conclude, improving catalytic efficiency by several orders of magnitude generally requires a long optimization trajectory. Progress is often hindered by diminishing returns and negative epistasis thus demanding increasingly more mutations to achieve diminishingly smaller improvements (Figure 2 suggests a rough estimate of 5 mutations per 10-fold for >1000-fold improvements). Successful navigation to the desired fitness peak is not guaranteed, but can be facilitated by boosting protein stability, by avoiding activity tradeoffs, by increasing screening or selection throughput and by exploring alternative genetic diversity (neutral drifts, InDels, or exchange of secondary structural elements). Combining directed evolution and computational design methods may also be fruitful – the latter allows to simultaneously introduce an entire set of mutations while the former complements the latter’s partial accuracy.

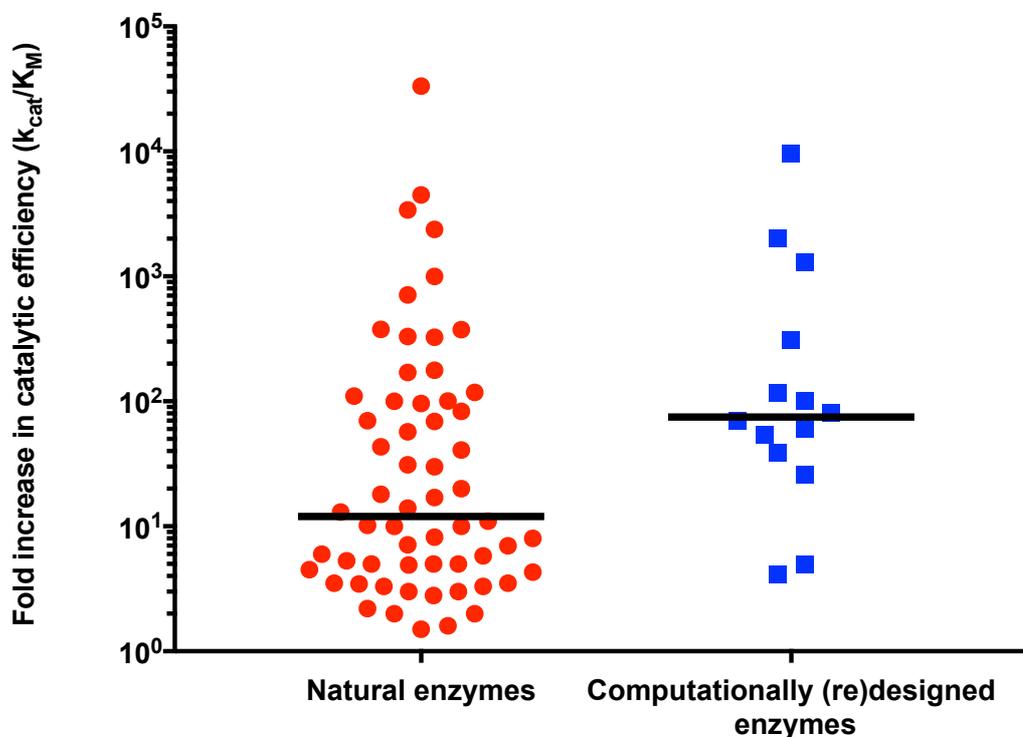


Figure 1. Laboratory optimizations of catalytic efficiency.

Shown are the fold-improvements in catalytic efficiency (k_{cat}/K_M) reported in the period of 2012-2016 for 60 natural enzymes optimized by directed evolution (red circles) [22,43,45,46,73-128], and for 14 enzymes that were computationally designed or redesigned and further optimized by directed evolution (2009-2016; blue squares) [7,44,50,129-139]. With few exceptions, the data describes the optimization of different enzymes. Indicated are the improvements for the most catalytically efficient variant compared to its starting point. Only studies that reported catalytic efficiency values of purified proteins, and described the number of directed evolution rounds and incorporated mutations, were included. The black, horizontal bars indicate the median fold improvement.

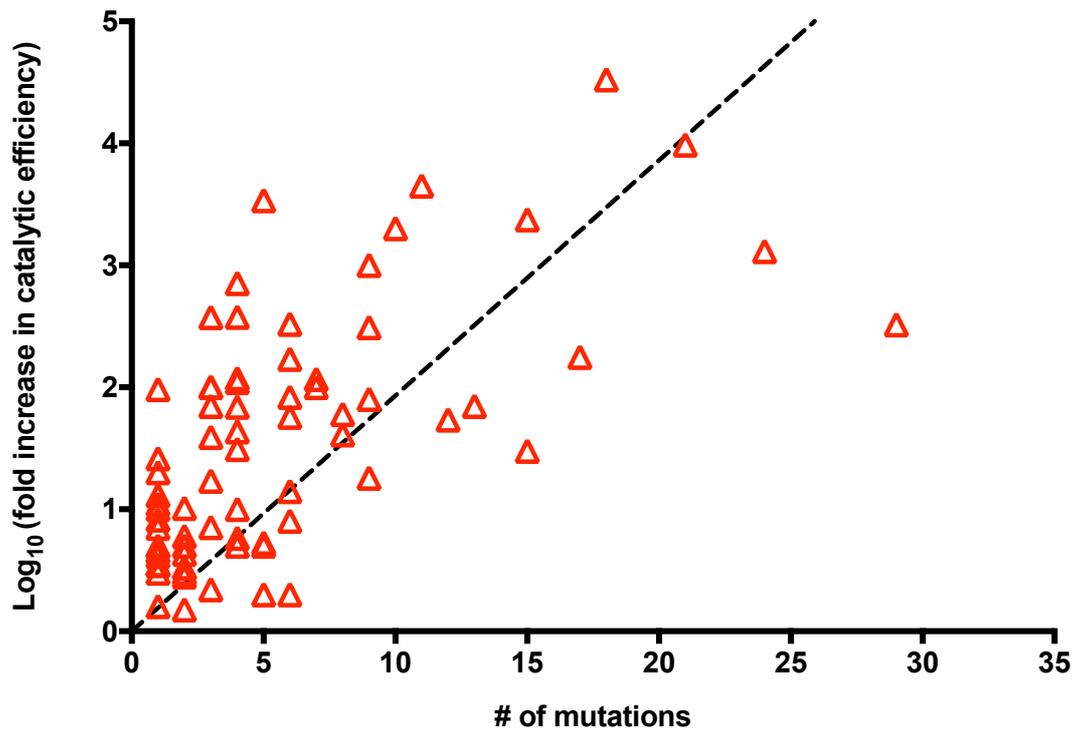


Figure 2. Large increases in catalytic efficiency demand multiple mutations. Plotted are the \log_{10} values of the fold-increases in k_{cat}/K_M of evolved enzymes *versus* the number of non-synonymous mutations in these variants (dataset of Figure 1). The line represents a linear fit ($y = 0.1931X \pm 0.01461$) suggesting that large improvements ($>10^3$ -fold) demand on average 5 mutations per order-of-magnitude improvement in catalytic efficiency.

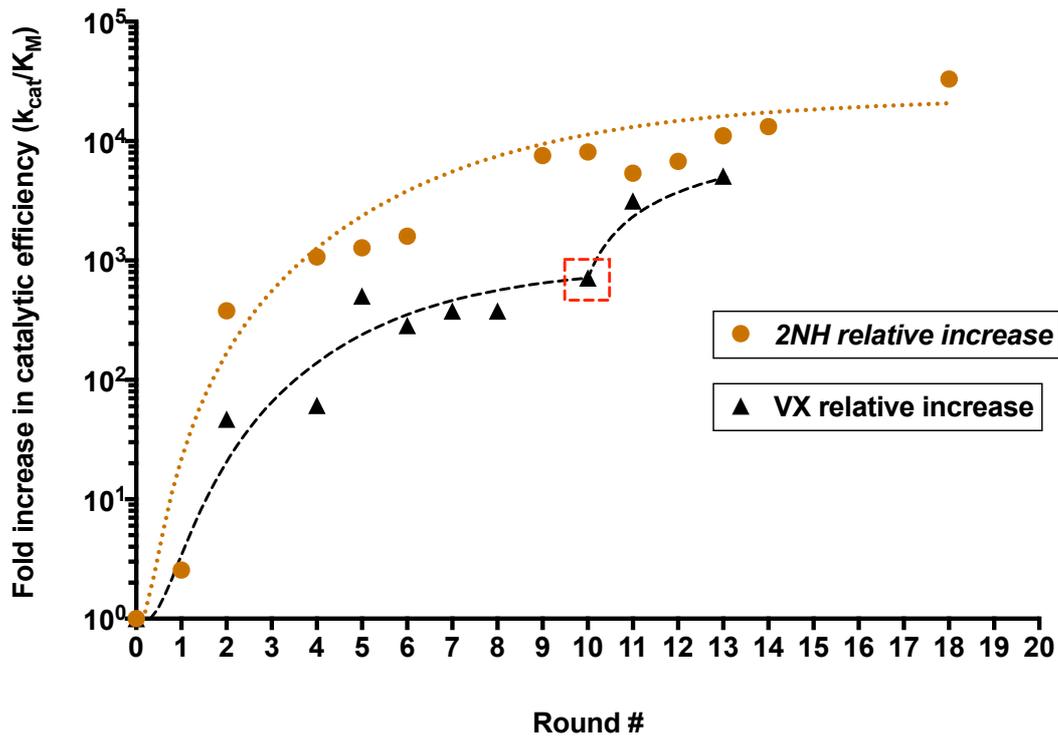
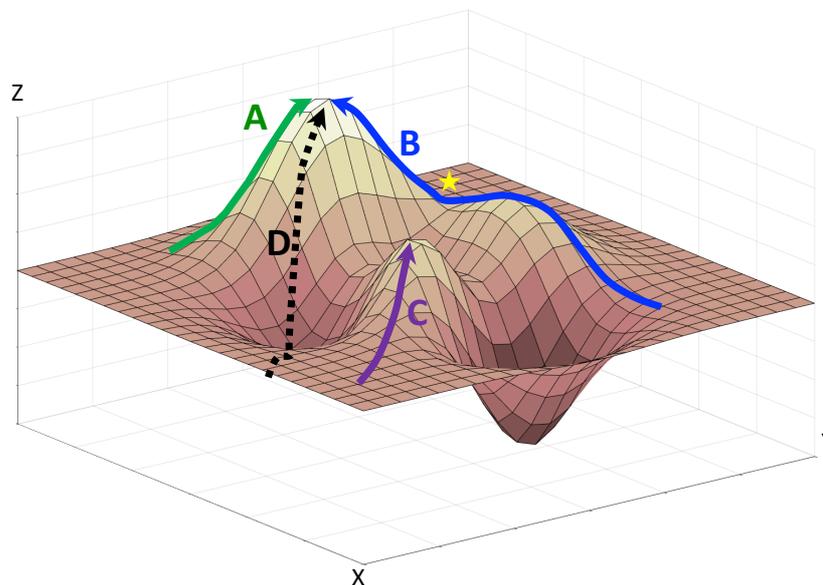


Figure 3. Enzyme optimizations are subject to diminishing returns. Plotted are the increases in catalytic efficiency observed in two different optimization trajectories of the same enzyme – a bacterial phosphotriesterase (PTE; Round 0 denotes the $k_{\text{cat}}/K_{\text{M}}$ value for wild-type with the target substrates). Optimization toward the hydrolysis of a synthetic aryl-ester substrate (2-naphthyl hexanoate) via directed evolution (●; data from [22]). Optimization toward hydrolysis of the toxic isomer of VX ($S_{\text{p}}\text{-VX}$) using directed evolution and computational design (▲; data from [24]). The red, dashed square denotes the introduction of a computationally designed variant containing 7 mutations that also improved $k_{\text{cat}}/K_{\text{M}}$ by ~2-fold. The lines depict the general trends in improvement for the two trajectories.

Box 1 – The landscape of evolutionary enzyme optimizations

In this schematic fitness landscape, the Z-axis indicates values of enzyme fitness that primarily relate to the enzyme's catalytic efficiency (k_{cat}/K_M), or catalytic turnover rate k_{cat} , under substrate saturation (although fitness also relates to protein folding, stability and solubility that dictate the amounts of soluble and active enzyme). The **fitness peak**, represented by the highest point on the Z axis, is defined per landscape; it represents the highest performance achievable for the trait under optimization given the explored sequence space. The limits of sequence exploration, schematically represented by the limits of the X-Y axes, are dictated by factors such as the number and type of mutations and the throughput of the applied screen/selection. If the landscape represents the entire sequence space, all possible sequences are included. Thus, the **global fitness peak** (*i.e.*, the maximal possible performance given the substrate and reaction) resides somewhere within the global sequence space (discussed in *How catalytically efficient can a given enzyme become?* in the main text).



The simplest optimization scenario is represented by a smooth (continuous) and gradual climb leading to the fitness peak (trajectory **A**). However, trajectories may be rugged as illustrated in trajectory **B** that reaches the peak through a **local plateau**

(yellow star). Such a trajectory is characterized by a key improving mutation/s that can manifest their beneficial fitness effect only at the background of other mutations that are either neutral or deleterious on their own (see a section on epistasis in main text).

Another scenario, represented by trajectory **C**, comprises an evolutionary 'dead end' as it leads to a **local peak** with much lower fitness than the fitness peak, and is separated from other trajectories by deep valleys of low fitness. Unlike the plateau depicted in trajectory **B**, the only way of traversing from the local peak in **C** to trajectories **A** or **B** is by reverting back to the starting point. Trajectories such as **C** are said to be *orthogonal* to others, as there is no way of crossing between them while avoiding a severe fitness loss. Orthogonal trajectories are usually initiated by different '**founder mutations**' as illustrated by the different locations of their starting points on the X-Y plane.

Trajectory **D** depicts a scenario in which the fitness peak is separated from the starting point by a deep, low-fitness valley. In this case, the only way to access it is by incorporating an entire set of mutations that mediate transition from the starting point to a new location in sequence space that can continuously lead to the fitness peak. This means that all possible subsets of these mutations would lead to loss of activity. The larger is the required mutational set, the more evolutionary inaccessible is an isolated peak.

Acknowledgments

Funding by a DTRA project grant (HDTRA1-11-C-0026), and by the European Commission H2020 FETOPEN grant project number: 686330 (Future Agriculture) are gratefully acknowledged.

References

1. Porter JL, Rusli RA, Ollis DL: **Directed Evolution of Enzymes for Industrial Biocatalysis**. *ChemBiochem* 2016, **17**:197-203.
2. Girvan HM, Munro AW: **Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology**. *Curr Opin Chem Biol* 2016, **31**:136-145.
3. Jemli S, Ayadi-Zouari D, Hlima HB, Bejar S: **Biocatalysts: application and engineering for industrial purposes**. *Crit Rev Biotechnol* 2016, **36**:246-258.
4. Songa EA, Okonkwo JO: **Recent approaches to improving selectivity and sensitivity of enzyme-based biosensors for organophosphorus pesticides: A review**. *Talanta* 2016, **155**:289-304.
5. Fuhrmann G, Leroux J-C: **Improving the Stability and Activity of Oral Therapeutic Enzymes—Recent Advances and Perspectives**. *Pharmaceutical Research* 2014, **31**:1099-1105.
6. Eisenthal R, Danson MJ, Hough DW: **Catalytic efficiency and k_{cat}/K_M : a useful comparator?** *Trends Biotechnol* 2007, **25**:247-249.
7. Obexer R, Pott M, Zeymer C, Griffiths AD, Hilvert D: **Efficient laboratory evolution of computationally designed enzymes with low starting activities using fluorescence-activated droplet sorting**. *Protein Eng Des Sel* 2016, **29**:355-366.
 - A high-throughput, high-accuracy enzyme screening technology. Although rather complex, its potential for pursuing long trajectories and obtaining large improvements is high.
8. Hancock SM, Rich JR, Caines ME, Strynadka NC, Withers SG: **Designer enzymes for glycosphingolipid synthesis by directed evolution**. *Nat Chem Biol* 2009, **5**:508-514.
9. Yew WS, Akana J, Wise EL, Rayment I, Gerlt JA: **Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase**. *Biochemistry* 2005, **44**:1807-1815.

10. Khare SD, Kipnis Y, Greisen P, Jr., Takeuchi R, Ashani Y, Goldsmith M, Song Y, Gallaher JL, Silman I, Leader H, et al.: **Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis.** *Nat Chem Biol* 2012, **8**:294-300.
11. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R: **The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters.** *Biochemistry* 2011, **50**:4402-4410.
12. Afriat L, Roodveldt C, Manco G, Tawfik DS: **The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase.** *Biochemistry* 2006, **45**:13677-13686.
13. Omburo GA, Kuo JM, Mullins LS, Raushel FM: **Characterization of the zinc binding site of bacterial phosphotriesterase.** *J Biol Chem* 1992, **267**:13278-13283.
14. Fasan R, Mehareenna YT, Snow CD, Poulos TL, Arnold FH: **Evolutionary history of a specialized p450 propane monooxygenase.** *J Mol Biol* 2008, **383**:1069-1080.
 - This paper described a long trajectory to evolve an enzymatic activity that is undetectable in the wild-type enzyme.
15. Rockah-Shmuel L, Tawfik DS: **Evolutionary transitions to new DNA methyltransferases through target site expansion and shrinkage.** *Nucleic Acids Res* 2012, **40**:11627-11637.
 - A series of long trajectories toward a range of new target specificities, including specificities that are undetectable in the wild-type enzyme.
16. Bar-Even A, Milo R, Noor E, Tawfik DS: **The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness.** *Biochemistry* 2015, **54**:4969-4977.
17. Reznik E, Mehta P, Segre D: **Flux imbalance analysis and the sensitivity of cellular growth to changes in metabolite pools.** *PLoS Comput Biol* 2013, **9**:e1003195.
18. Sajitz-Hermstein M, Nikoloski Z: **Multi-objective shadow prices point at principles of metabolic regulation.** *Biosystems* 2016, **146**:91-101.
19. Bommarius AS, Paye MF: **Stabilizing biocatalysts.** *Chem Soc Rev* 2013, **42**:6534-6565.
20. Socha RD, Tokuriki N: **Modulating protein stability - directed evolution strategies for improved protein function.** *FEBS J* 2013, **280**:5582-5595.
21. Eijsink VG, Gaseidnes S, Borchert TV, van den Burg B: **Directed evolution of enzyme stability.** *Biomol Eng* 2005, **22**:21-30.

22. Tokuriki N, Jackson CJ, Afriat-Jurnou L, Wyganowski KT, Tang R, Tawfik DS: **Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme.** *Nat Commun* 2012, **3**:1257.
- An example of an 18-round long directed evolution experiment, in which the activity of an enzyme towards a new substrate was greatly increased in spite of strong diminishing returns.
23. Campbell E, Kaltenbach M, Correy GJ, Carr PD, Porebski BT, Livingstone EK, Afriat-Jurnou L, Buckle AM, Weik M, Hollfelder F, et al.: **The role of protein dynamics in the evolution of new enzyme function.** *Nat Chem Biol* 2016, **12**:944-950.
24. Goldsmith M, Aggarwal N, Ashani Y, Jubran H, Greisen PJ, Ovchinnikov S, Leader H, Baker D, Sussman JL, Goldenzweig A, et al.: **Overcoming an optimization plateau in the directed evolution of highly efficient nerve agent bioscavengers.** *Protein Eng Des Sel* 2017, 10.1093/protein/gzx003:1-13.
- A recent description of enzyme optimization for an in-vivo application. Optimization plateaus and the means for overcoming them are described.
25. Goldenzweig A, Goldsmith M, Hill SE, Gertman O, Laurino P, Ashani Y, Dym O, Unger T, Albeck S, Prilusky J, et al.: **Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability.** *Mol Cell* 2016, **63**:337-346.
- A computational design method for boosting protein stability that can also be applied to boost enzyme evolvability and to overcome optimization plateaus.
26. Yip SH, Matsumura I: **Substrate ambiguous enzymes within the Escherichia coli proteome offer different evolutionary solutions to the same problem.** *Mol Biol Evol* 2013, **30**:2001-2012.
27. Porter JL, Collyer CA, Ollis DL: **Compensatory stabilizing role of surface mutations during the directed evolution of diene lactone hydrolase for enhanced activity.** *Protein J* 2015, **34**:82-89.
28. Petrie KL, Joyce GF: **Limits of neutral drift: lessons from the in vitro evolution of two ribozymes.** *J Mol Evol* 2014, **79**:75-90.
29. Toth-Petroczy A, Tawfik DS: **Hopeful (protein InDel) monsters?** *Structure* 2014, **22**:803-804.
30. Jones DD, Arpino JA, Baldwin AJ, Edmundson MC: **Transposon-based approaches for generating novel molecular diversity during directed evolution.** *Methods Mol Biol* 2014, **1179**:159-172.
31. Kipnis Y, Dellus-Gur E, Tawfik DS: **TRINS: a method for gene modification by randomized tandem repeat insertions.** *Protein Eng Des Sel* 2012, **25**:437-444.

32. Smith MA, Arnold FH: **Designing libraries of chimeric proteins using SCHEMA recombination and RASPP.** *Methods Mol Biol* 2014, **1179**:335-343.
- This work describes two algorithms for designing libraries based on the exchange of whole secondary structural elements between similar proteins using homologous recombination.
33. Miton CM, Tokuriki N: **How mutational epistasis impairs predictability in protein evolution and design.** *Protein Sci* 2016, **25**:1260-1272.
34. Starr TN, Thornton JW: **Epistasis in protein evolution.** *Protein Sci* 2016, **25**:1204-1218.
35. Dellus-Gur E, Elias M, Caselli E, Prati F, Salverda ML, de Visser JA, Fraser JS, Tawfik DS: **Negative Epistasis and Evolvability in TEM-1 beta-Lactamase--The Thin Line between an Enzyme's Conformational Freedom and Disorder.** *J Mol Biol* 2015, **427**:2396-2409.
36. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, St Clair JL, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL, et al.: **Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction.** *Science* 2010, **329**:309-313.
37. Ovchinnikov S, Park H, Varghese N, Huang PS, Pavlopoulos GA, Kim DE, Kamisetty H, Kyrpides NC, Baker D: **Protein structure determination using metagenome sequence data.** *Science* 2017, **355**:294-298.
38. Strafford J, Payongsri P, Hibbert EG, Morris P, Batth SS, Steadman D, Smith ME, Ward JM, Hailes HC, Dalby PA: **Directed evolution to re-adapt a co-evolved network within an enzyme.** *J Biotechnol* 2012, **157**:237-245.
39. Sullivan BJ, Durani V, Magliery TJ: **Triosephosphate isomerase by consensus design: dramatic differences in physical properties and activity of related variants.** *J Mol Biol* 2011, **413**:195-208.
40. Behrendorff JB, Huang W, Gillam EM: **Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity.** *Biochem J* 2015, **467**:1-15.
41. Denard CA, Ren H, Zhao H: **Improving and repurposing biocatalysts via directed evolution.** *Curr Opin Chem Biol* 2015, **25**:55-64.
42. Kataoka M, Miyakawa T, Shimizu S, Tanokura M: **Enzymes useful for chiral compound synthesis: structural biology, directed evolution, and protein engineering for industrial use.** *Appl Microbiol Biotechnol* 2016, **100**:5747-5757.

43. Alvizo O, Nguyen LJ, Savile CK, Bresson JA, Lakhapatri SL, Solis EO, Fox RJ, Broering JM, Benoit MR, Zimmerman SA, et al.: **Directed evolution of an ultrastable carbonic anhydrase for highly efficient carbon capture from flue gas.** *Proc Natl Acad Sci U S A* 2014, **111**:16436-16441.
- An enzyme optimization effort aimed at achieving high performance for an industrial application.
44. Althoff EA, Wang L, Jiang L, Giger L, Lassila JK, Wang Z, Smith M, Hari S, Kast P, Herschlag D, et al.: **Robust design and optimization of retroaldol enzymes.** *Protein Sci* 2012, **21**:717-726.
45. Bjelic S, Kipnis Y, Wang L, Pianowski Z, Vorobiev S, Su M, Seetharaman J, Xiao R, Kornhaber G, Hunt JF, et al.: **Exploration of alternate catalytic mechanisms and optimization strategies for retroaldolase design.** *J Mol Biol* 2014, **426**:256-271.
46. Giger L, Caner S, Obexer R, Kast P, Baker D, Ban N, Hilvert D: **Evolution of a designed retro-aldolase leads to complete active site remodeling.** *Nat Chem Biol* 2013, **9**:494-498.
- This work describes a long evolutionary trajectory aiming at obtain high catalytic efficiency in a computationally designed enzyme with an initial low activity.
47. Griswold KE, Aiyappan NS, Iverson BL, Georgiou G: **The evolution of catalytic efficiency and substrate promiscuity in human theta class 1-1 glutathione transferase.** *J Mol Biol* 2006, **364**:400-410.
48. Obexer R, Godina A, Garrabou X, Mittl PR, Baker D, Griffiths AD, Hilvert D: **Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase.** *Nat Chem* 2017, **9**:50-56.
- The application of a high-throughput, high-accuracy enzyme screening technology (Ref. 7) to obtain a large improvement in catalytic efficiency.
49. Oue S, Okamoto A, Yano T, Kagamiyama H: **Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues.** *J Biol Chem* 1999, **274**:2344-2349.
50. Preiswerk N, Beck T, Schulz JD, Milovnik P, Mayer C, Siegel JB, Baker D, Hilvert D: **Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase.** *Proc Natl Acad Sci U S A* 2014, **111**:8013-8018.
51. Tokuriki N, Tawfik DS: **Stability effects of mutations and protein evolvability.** *Curr Opin Struct Biol* 2009, **19**:596-604.
52. Cox VE, Gaucher EA: **Engineering proteins by reconstructing evolutionary adaptive paths.** *Methods Mol Biol* 2014, **1179**:353-363.

53. Zakas PM, Brown HC, Knight K, Meeks SL, Spencer HT, Gaucher EA, Doering CB: **Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction.** *Nat Biotechnol* 2017, **35**:35-37.
54. Romero-Romero ML, Risso VA, Martinez-Rodriguez S, Ibarra-Molero B, Sanchez-Ruiz JM: **Engineering ancestral protein hyperstability.** *Biochem J* 2016, **473**:3611-3620.
55. Trudeau DL, Kaltenbach M, Tawfik DS: **On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins.** *Mol Biol Evol* 2016, **33**:2633-2641.
56. Goldsmith M, Tawfik DS: **Enzyme engineering by targeted libraries.** *Methods Enzymol* 2013, **523**:257-283.
57. Zhang L, Tang X, Cui D, Yao Z, Gao B, Jiang S, Yin B, Yuan YA, Wei D: **A method to rationally increase protein stability based on the charge-charge interaction, with application to lipase LipK107.** *Protein Sci* 2014, **23**:110-116.
58. Bednar D, Beerens K, Sebestova E, Bendl J, Khare S, Chaloupkova R, Prokop Z, Brezovsky J, Baker D, Damborsky J: **FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants.** *PLoS Comput Biol* 2015, **11**:e1004556.
 - A computational method based on energy and phylogeny considerations that can be used to stabilize proteins.
59. Durani V, Magliery TJ: **Protein engineering and stabilization from sequence statistics: variation and covariation analysis.** *Methods Enzymol* 2013, **523**:237-256.
60. Dellus-Gur E, Toth-Petroczy A, Elias M, Tawfik DS: **What makes a protein fold amenable to functional innovation? Fold polarity and stability trade-offs.** *J Mol Biol* 2013, **425**:2609-2621.
61. Leemhuis H, Kelly RM, Dijkhuizen L: **Directed evolution of enzymes: Library screening strategies.** *IUBMB Life* 2009, **61**:222-228.
62. Hubbard BP, Badran AH, Zuris JA, Guilinger JP, Davis KM, Chen L, Tsai SQ, Sander JD, Joung JK, Liu DR: **Continuous directed evolution of DNA-binding proteins to improve TALEN specificity.** *Nat Methods* 2015, **12**:939-942.
63. Pu J, Zinkus-Boltz J, Dickinson BC: **Evolution of a split RNA polymerase as a versatile biosensor platform.** *Nat Chem Biol* 2017, **13**:432-438.
64. Carlson JC, Badran AH, Guggiana-Nilo DA, Liu DR: **Negative selection and stringency modulation in phage-assisted continuous evolution.** *Nat Chem Biol* 2014, **10**:216-222.

65. Badran AH, Guzov VM, Huai Q, Kemp MM, Vishwanath P, Kain W, Nance AM, Evdokimov A, Moshiri F, Turner KH, et al.: **Continuous evolution of *Bacillus thuringiensis* toxins overcomes insect resistance.** *Nature* 2016, **533**:58-63.
66. Gielen F, Hours R, Emond S, Fischlechner M, Schell U, Hollfelder F: **Ultrahigh-throughput-directed enzyme evolution by absorbance-activated droplet sorting (AADS).** *Proc Natl Acad Sci U S A* 2016, **113**:E7383-E7389.
- A high-throughput, high-accuracy enzyme screening technology that can be applied for pursuing long trajectories.
67. Amitai G, Gupta RD, Tawfik DS: **Latent evolutionary potentials under the neutral mutational drift of an enzyme.** *HFSP J* 2007, **1**:67-78.
68. Bershtein S, Goldin K, Tawfik DS: **Intense neutral drifts yield robust and evolvable consensus proteins.** *J Mol Biol* 2008, **379**:1029-1044.
69. Bershtein S, Tawfik DS: **Ohno's model revisited: measuring the frequency of potentially adaptive mutations under various mutational drifts.** *Mol Biol Evol* 2008, **25**:2311-2318.
70. Gupta RD, Tawfik DS: **Directed enzyme evolution via small and effective neutral drift libraries.** *Nat Methods* 2008, **5**:939-942.
71. Smith WS, Hale JR, Neylon C: **Applying neutral drift to the directed molecular evolution of a beta-glucuronidase into a beta-galactosidase: Two different evolutionary pathways lead to the same variant.** *BMC Res Notes* 2011, **4**:138.
72. Steinberg B, Ostermeier M: **Environmental changes bridge evolutionary valleys.** *Sci Adv* 2016, **2**:e1500921.
73. Li Z, Roccatano D, Lorenz M, Schwaneberg U: **Directed evolution of subtilisin E into a highly active and guanidinium chloride- and sodium dodecylsulfate-tolerant protease.** *Chembiochem* 2012, **13**:691-699.
74. Goldsmith M, Ashani Y, Simo Y, Ben-David M, Leader H, Silman I, Sussman JL, Tawfik DS: **Evolved stereoselective hydrolases for broad-spectrum G-type nerve agent detoxification.** *Chem Biol* 2012, **19**:456-466.
75. Jeong YS, Choi SL, Kyeong HH, Kim JH, Kim EJ, Pan JG, Rha E, Song JJ, Lee SG, Kim HS: **High-throughput screening system based on phenolics-responsive transcription activator for directed evolution of organophosphate-degrading enzymes.** *Protein Eng Des Sel* 2012, **25**:725-731.
76. Kintsjes B, Hein C, Mohamed MF, Fischlechner M, Courtois F, Laine C, Hollfelder F: **Picoliter cell lysate assays in microfluidic droplet**

- compartments for directed enzyme evolution.** *Chem Biol* 2012, **19**:1001-1009.
77. Smith MA, Romero PA, Wu T, Brustad EM, Arnold FH: **Chimeragenesis of distantly-related proteins by noncontiguous recombination.** *Protein Sci* 2013, **22**:231-238.
78. Brinkmann-Chen S, Flock T, Cahn JK, Snow CD, Brustad EM, McIntosh JA, Meinhold P, Zhang L, Arnold FH: **General approach to reversing ketol-acid reductoisomerase cofactor dependence from NADPH to NADH.** *Proc Natl Acad Sci U S A* 2013, **110**:10946-10951.
79. Blomberg R, Kries H, Pinkas DM, Mittl PR, Grutter MG, Privett HK, Mayo SL, Hilvert D: **Precision is essential for efficient catalysis in an evolved Kemp eliminase.** *Nature* 2013, **503**:418-421.
- This work describes a long directed evolution effort to enhance the activity of a computationally designed enzyme.
80. Muller MM, Allison JR, Hongdilokkul N, Gaillon L, Kast P, van Gunsteren WF, Marliere P, Hilvert D: **Directed evolution of a model primordial enzyme provides insights into the development of the genetic code.** *PLoS Genet* 2013, **9**:e1003187.
81. Wellner A, Raitses Gurevich M, Tawfik DS: **Mechanisms of protein sequence divergence and incompatibility.** *PLoS Genet* 2013, **9**:e1003665.
82. Pengpai Z, Sheng H, Lehe M, Yinlin L, Zhihua J, Guixiang H: **Improving the activity of cytochrome P450 BM-3 catalyzing indole hydroxylation by directed evolution.** *Appl Biochem Biotechnol* 2013, **171**:93-103.
83. Sideri A, Goyal A, Di Nardo G, Tsotsou GE, Gilardi G: **Hydroxylation of non-substituted polycyclic aromatic hydrocarbons by cytochrome P450 BM3 engineered by directed evolution.** *J Inorg Biochem* 2013, **120**:1-7.
84. Tan LT, Hiraishi T, Sudesh K, Maeda M: **Directed evolution of poly[(R)-3-hydroxybutyrate] depolymerase using cell surface display system: functional importance of asparagine at position 285.** *Appl Microbiol Biotechnol* 2013, **97**:4859-4871.
85. Wang J, Zhang Q, Huang Z, Liu Z: **Directed evolution of a family 26 glycoside hydrolase: endo-beta-1, 4-mannanase from *Pantoea agglomerans* A021.** *J Biotechnol* 2013, **167**:350-356.
86. Wang Y, Feng S, Zhan T, Huang Z, Wu G, Liu Z: **Improving catalytic efficiency of endo-beta-1, 4-xylanase from *Geobacillus stearothermophilus* by directed evolution and H179 saturation mutagenesis.** *J Biotechnol* 2013, **168**:341-347.

87. Choi S, Han S, Lee H, Chun YJ, Kim D: **Evaluation of Luminescent P450 Analysis for Directed Evolution of Human CYP4A11.** *Biomol Ther (Seoul)* 2013, **21**:487-492.
88. Wang Y, Li X, Chen X, Chen D: **Directed evolution and characterization of atrazine chlorohydrolase variants with enhanced activity.** *Biochemistry (Mosc)* 2013, **78**:1104-1111.
89. Bernath-Levin K, Shainsky J, Sigawi L, Fishman A: **Directed evolution of nitrobenzene dioxygenase for the synthesis of the antioxidant hydroxytyrosol.** *Appl Microbiol Biotechnol* 2014, **98**:4975-4985.
90. Giacobbe S, Vincent F, Faraco V: **Development of an improved variant of GH51 alpha-l-arabinofuranosidase from *Pleurotus ostreatus* by directed evolution.** *N Biotechnol* 2014, **31**:230-236.
91. Liu M, Xie W, Xu H, Gu J, Lv X, Yu H, Ye L: **Directed evolution of an exoglucanase facilitated by a co-expressed beta-glucosidase and construction of a whole engineered cellulase system in *Escherichia coli*.** *Biotechnol Lett* 2014, **36**:1801-1807.
92. Lu WC, Levy M, Kincaid R, Ellington AD: **Directed evolution of the substrate specificity of biotin ligase.** *Biotechnol Bioeng* 2014, **111**:1071-1081.
93. Molina-Espeja P, Garcia-Ruiz E, Gonzalez-Perez D, Ullrich R, Hofrichter M, Alcalde M: **Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*.** *Appl Environ Microbiol* 2014, **80**:3496-3507.
94. Ostafe R, Prodanovic R, Nazor J, Fischer R: **Ultra-high-throughput screening method for the directed evolution of glucose oxidase.** *Chem Biol* 2014, **21**:414-421.
95. Shi F, Xie Y, Jiang J, Wang N, Li Y, Wang X: **Directed evolution and mutagenesis of glutamate decarboxylase from *Lactobacillus brevis* Lb85 to broaden the range of its activity toward a near-neutral pH.** *Enzyme Microb Technol* 2014, **61-62**:35-43.
96. Wang NQ, Sun J, Huang J, Wang P: **Cloning, expression, and directed evolution of carbonyl reductase from *Leifsonia xyli* HS0904 with enhanced catalytic efficiency.** *Appl Microbiol Biotechnol* 2014, **98**:8591-8601.
97. Zhang ZG, Lonsdale R, Sanchis J, Reetz MT: **Extreme synergistic mutational effects in the directed evolution of a baeyer-villiger monooxygenase as catalyst for asymmetric sulfoxidation.** *J Am Chem Soc* 2014, **136**:17262-17272.
98. Jimenez-Oses G, Osuna S, Gao X, Sawaya MR, Gilson L, Collier SJ, Huisman GW, Yeates TO, Tang Y, Houk KN: **The role of distant mutations and allosteric regulation on LovD active site dynamics.** *Nat Chem Biol* 2014, **10**:431-436.

99. Schulenburg C, Stark Y, Kunzle M, Hilvert D: **Comparative laboratory evolution of ordered and disordered enzymes.** *J Biol Chem* 2015, **290**:9310-9320.
100. Bosshart A, Hee CS, Bechtold M, Schirmer T, Panke S: **Directed divergent evolution of a thermostable D-tagatose epimerase towards improved activity for two hexose substrates.** *Chembiochem* 2015, **16**:592-601.
101. Buller AR, Brinkmann-Chen S, Romney DK, Herger M, Murciano-Calles J, Arnold FH: **Directed evolution of the tryptophan synthase beta-subunit for stand-alone function recapitulates allosteric activation.** *Proc Natl Acad Sci U S A* 2015, **112**:14599-14604.
102. Cheng F, Zhu L, Lue H, Bernhagen J, Schwaneberg U: **Directed arginine deiminase evolution for efficient inhibition of arginine-auxotrophic melanomas.** *Appl Microbiol Biotechnol* 2015, **99**:1237-1247.
103. Han SS, Park WJ, Kim HS, Kim GJ: **Antisense RNA-based High-Throughput Screen System for Directed Evolution of Quorum Quenching Enzymes.** *ACS Chem Biol* 2015, **10**:2598-2605.
104. Jiang W, Zhuang Y, Wang S, Fang B: **Directed Evolution and Resolution Mechanism of 1, 3-Propanediol Oxidoreductase from *Klebsiella pneumoniae* toward Higher Activity by Error-Prone PCR and Bioinformatics.** *PLoS One* 2015, **10**:e0141837.
105. Krause M, Neubauer P, Wierenga RK: **Structure-based directed evolution of a monomeric triosephosphate isomerase: toward a pentose sugar isomerase.** *Protein Eng Des Sel* 2015, **28**:187-197.
106. Kwan DH, Constantinescu I, Chapanian R, Higgins MA, Kotzler MP, Samain E, Boraston AB, Kizhakkedathu JN, Withers SG: **Toward Efficient Enzymes for the Generation of Universal Blood through Structure-Guided Directed Evolution.** *J Am Chem Soc* 2015, **137**:5695-5705.
107. Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, Mootha VK, Ting AY: **Directed evolution of APEX2 for electron microscopy and proximity labeling.** *Nat Methods* 2015, **12**:51-54.
108. Lulsdorf N, Pitzler C, Biggel M, Martinez R, Vojcic L, Schwaneberg U: **A flow cytometer-based whole cell screening toolbox for directed hydrolase evolution through fluorescent hydrogels.** *Chem Commun (Camb)* 2015, **51**:8679-8682.
109. Ng TK, Gahan LR, Schenk G, Ollis DL: **Altering the substrate specificity of methyl parathion hydrolase with directed evolution.** *Arch Biochem Biophys* 2015, **573**:59-68.

110. Porter JL, Boon PL, Murray TP, Huber T, Collyer CA, Ollis DL: **Directed evolution of new and improved enzyme functions using an evolutionary intermediate and multidirectional search.** *ACS Chem Biol* 2015, **10**:611-621.
- A 14 round-long directed evolution effort to enhanced both activity and stability of the target enzyme.
111. Taylor JL, Price JE, Toney MD: **Directed evolution of the substrate specificity of dialkylglycine decarboxylase.** *Biochim Biophys Acta* 2015, **1854**:146-155.
112. Vina-Gonzalez J, Gonzalez-Perez D, Ferreira P, Martinez AT, Alcalde M: **Focused Directed Evolution of Aryl-Alcohol Oxidase in *Saccharomyces cerevisiae* by Using Chimeric Signal Peptides.** *Appl Environ Microbiol* 2015, **81**:6451-6462.
113. Wan NW, Liu ZQ, Xue F, Huang K, Tang LJ, Zheng YG: **An efficient high-throughput screening assay for rapid directed evolution of halohydrin dehalogenase for preparation of beta-substituted alcohols.** *Appl Microbiol Biotechnol* 2015, **99**:4019-4029.
114. Herger M, van Roye P, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH: **Synthesis of beta-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase.** *J Am Chem Soc* 2016, **138**:8388-8391.
115. Cahn JK, Baumschlager A, Brinkmann-Chen S, Arnold FH: **Mutations in adenine-binding pockets enhance catalytic properties of NAD(P)H-dependent enzymes.** *Protein Eng Des Sel* 2016, **29**:31-38.
116. Garrabou X, Wicky BI, Hilvert D: **Fast Knoevenagel Condensations Catalyzed by an Artificial Schiff-Base-Forming Enzyme.** *J Am Chem Soc* 2016, **138**:6972-6974.
117. Cheng Q, Gao H, Hu N: **A trehalase from *Zunongwangia* sp.: characterization and improving catalytic efficiency by directed evolution.** *BMC Biotechnol* 2016, **16**:9.
118. Kim JE, Huang R, Chen H, You C, Zhang YH: **Facile Construction of Random Gene Mutagenesis Library for Directed Evolution Without the Use of Restriction Enzyme in *Escherichia coli*.** *Biotechnol J* 2016, **11**:1142-1150.
119. Jeschek M, Reuter R, Heinisch T, Trindler C, Klehr J, Panke S, Ward TR: **Directed evolution of artificial metalloenzymes for in vivo metathesis.** *Nature* 2016, **537**:661-665.
120. Kan SB, Lewis RD, Chen K, Arnold FH: **Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life.** *Science* 2016, **354**:1048-1051.

121. Luo H, Ma J, Chang Y, Yu H, Shen Z: **Directed Evolution and Mutant Characterization of Nitrilase from *Rhodococcus rhodochrous* tg1-A6.** *Appl Biochem Biotechnol* 2016, **178**:1510-1521.
122. Molina-Espeja P, Canellas M, Plou FJ, Hofrichter M, Lucas F, Guallar V, Alcalde M: **Synthesis of 1-Naphthol by a Natural Peroxygenase Engineered by Directed Evolution.** *Chembiochem* 2016, **17**:341-349.
123. Sun H, Yeo WL, Lim YH, Chew X, Smith DJ, Xue B, Chan KP, Robinson RC, Robins EG, Zhao H, et al.: **Directed Evolution of a Fluorinase for Improved Fluorination Efficiency with a Non-native Substrate.** *Angew Chem Int Ed Engl* 2016, **55**:14277-14280.
124. Xu X, Liu MQ, Huo WK, Dai XJ: **Obtaining a mutant of *Bacillus amyloliquefaciens* xylanase A with improved catalytic activity by directed evolution.** *Enzyme Microb Technol* 2016, **86**:59-66.
125. Zhang K, Guo Y, Yao P, Lin Y, Kumar A, Liu Z, Wu G, Zhang L: **Characterization and directed evolution of BliGO, a novel glycine oxidase from *Bacillus licheniformis*.** *Enzyme Microb Technol* 2016, **85**:12-18.
126. Zhu L, Wu Z, Jin JM, Tang SY: **Directed evolution of leucine dehydrogenase for improved efficiency of L-tert-leucine synthesis.** *Appl Microbiol Biotechnol* 2016, **100**:5805-5813.
127. Zhu X, Shieh P, Su M, Bertozzi CR, Zhang W: **A fluorogenic screening platform enables directed evolution of an alkyne biosynthetic tool.** *Chem Commun (Camb)* 2016, **52**:11239-11242.
128. Goldsmith M, Ashani Y, Margalit R, Nyska A, Mirelman D, Tawfik DS: **A new post-intoxication treatment of paraoxon and parathion poisonings using an evolved PON1 variant and recombinant GOT1.** *Chem Biol Interact* 2016, **259**:242-251.
129. Barajas JF, Phelan RM, Schaub AJ, Kliewer JT, Kelly PJ, Jackson DR, Luo R, Keasling JD, Tsai SC: **Comprehensive Structural and Biochemical Analysis of the Terminal Myxalamid Reductase Domain for the Engineered Production of Primary Alcohols.** *Chem Biol* 2015, **22**:1018-1029.
130. Bjelic S, Nivon LG, Celebi-Olcum N, Kiss G, Rosewall CF, Lovick HM, Ingalls EL, Gallaher JL, Seetharaman J, Lew S, et al.: **Computational design of enone-binding proteins with catalytic activity for the Morita-Baylis-Hillman reaction.** *ACS Chem Biol* 2013, **8**:749-757.
131. Chen CY, Georgiev I, Anderson AC, Donald BR: **Computational structure-based redesign of enzyme activity.** *Proc Natl Acad Sci U S A* 2009, **106**:3764-3769.

132. Gordon SR, Stanley EJ, Wolf S, Toland A, Wu SJ, Hadidi D, Mills JH, Baker D, Pultz IS, Siegel JB: **Computational design of an alpha-gliadin peptidase**. *J Am Chem Soc* 2012, **134**:20513-20520.
133. Khersonsky O, Kiss G, Rothlisberger D, Dym O, Albeck S, Houk KN, Baker D, Tawfik DS: **Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59**. *Proc Natl Acad Sci U S A* 2012, **109**:10358-10363.
134. Liu L, Murphy P, Baker D, Lutz S: **Computational design of orthogonal nucleoside kinases**. *Chem Commun (Camb)* 2010, **46**:8803-8805.
135. Murphy PM, Bolduc JM, Gallaher JL, Stoddard BL, Baker D: **Alteration of enzyme specificity by computational loop remodeling and design**. *Proc Natl Acad Sci U S A* 2009, **106**:9215-9220.
136. Nannemann DP, Kaufmann KW, Meiler J, Bachmann BO: **Design and directed evolution of a dideoxy purine nucleoside phosphorylase**. *Protein Eng Des Sel* 2010, **23**:607-616.
137. Richter F, Blomberg R, Khare SD, Kiss G, Kuzin AP, Smith AJ, Gallaher J, Pianowski Z, Helgeson RC, Grjasnow A, et al.: **Computational design of catalytic dyads and oxyanion holes for ester hydrolysis**. *J Am Chem Soc* 2012, **134**:16197-16206.
138. Sirin S, Kumar R, Martinez C, Karmilowicz MJ, Ghosh P, Abramov YA, Martin V, Sherman W: **A computational approach to enzyme design: predicting omega-aminotransferase catalytic activity using docking and MM-GBSA scoring**. *J Chem Inf Model* 2014, **54**:2334-2346.
139. Zhang W, Moden O, Tars K, Mannervik B: **Structure-based redesign of GST A2-2 for enhanced catalytic efficiency with azathioprine**. *Chem Biol* 2012, **19**:414-421.