Research Article

Model and Laboratory Demonstrations That Evolutionary Optimization Works Well Only If Preceded by Invention—Selection Itself Is Not Inventive

Douglas D. Axe* and Ann K. Gauger

Biologic Institute, Redmond, Washington, USA

Abstract

Since biological inventions only benefit their possessors after they work, their origins cannot be attributed to their selective effects. One proposed solution to this conundrum is that selection perfects activities that already existed in rudimentary form before they became beneficial. An example of this idea for protein origins is the *promiscuity* hypothesis, which claims that minor aberrant side-reactions in enzymes can be evolutionary starting points for proficient new enzymes. Another example—the *junk* hypothesis—claims that proteins arising from accidental expression of non-genic DNA may likewise have slight activities that, through evolutionary optimization, lead to proficient enzymes. Here, we tested these proposals by observing how the endpoint of simple evolutionary optimization depends on the starting point. Beginning with optimization of protein-like constructs in the *Stylus* computational model, we compared promiscuous and junk starting points, where design elements specific to the test function were completely absent, to a starting point that retained most elements of a good design (mutation having disrupted some). In all three cases, evolutionary optimization improved activities by a large factor. The extreme weakness of the original activities, however, meant even large improvements could be inconsequential. Indeed, the endpoint was itself a proficient design only in the case where this design was largely present from the outset. Laboratory optimization of ampicillin-resistance proteins derived from a natural β-lactamase produced similar results. Our junk protein here was a deletion mutant that somehow confers weak resistance without the original catalytic mechanism (much of the active site having been lost). Evolutionary optimization was unable to improve that mutant. In contrast, a comparably weak mutant that retained the active site surpassed the natural β-lactamase after six rounds of selection. So, while mutation and selection can improve the proficiency of good designs through small structural adjustments, they seem unable to convert fortuitous selectable activities into good designs.

Cite as: Axe DD, Gauger AK (2015) Model and laboratory demonstrations that evolutionary optimization works well only if preceded by invention—Selection itself is not inventive. BIO-Complexity 2015 (2):1–13. **doi:**10.5048/BIO-C.2015.2.

Editor: Robert J. Marks II

Received: September 23, 2015; **Accepted:** November 26, 2015; **Published:** December 22, 2015

Note: A *Critique* of this paper, when available, will be assigned **doi:**10.5048/BIO-C.2015.2.c.

*Email: daxe@biologicinstitute.org

INTRODUCTION

Over twenty years ago, in a paper titled *"The arrival of the fittest": Toward a theory of biological organization*, Fontana and Buss pointed out that the modern conception of evolution "suffers from an existence problem" [1]. By this they meant that "present theory tacitly assumes the prior existence of the entities whose features it is meant to explain" [1]. Indeed, the logic of natural selection has raised this existence problem from the very beginning of Darwin's theory. Under favorable circumstances, natural selection causes the genes carried by the most fit members of a species to be passed on to future generations. But this only explains which existing variations are kept and which are lost, and since these variations are never seen to include completely new functional features (to say nothing of new organismal kinds), an honest assessment of the facts forces us to conclude that the production of new forms of life is not actually explained by current evolutionary theory.

What Fontana and Buss called the existence problem we will refer to as *the problem of invention*, where by "invention" we simply mean the process by which any remarkable functional thing first came to exist. Much hangs on the solution to this problem, whatever we choose to call it. Dutch botanist Hugo De Vries pointed to the explanatory hole in Darwin's theory in 1904, quoting an acquaintance of his as saying that "Natural selection may explain the *survival* of the fittest, but it cannot

explain the *arrival* of the fittest" [2]. By referring to this quote in their title, Fontana and Buss emphasized that this hole has been long recognized, and while they hoped their theory of biological organization would fill it, the truth is that it remains as prominent today as ever. Twenty years after their paper, the title of a book by Swiss evolutionary biologist Andreas Wagner repeats the phrase that has now epitomized the problem for over a century: *Arrival of the Fittest—Solving Evolution's Greatest Puzzle*. As Wagner puts it, "Natural selection can *preserve* innovations, but, it cannot *create* them" [3].

Running counter to this long-standing critique of natural selection has been the equally long tradition of claiming, as Darwin did, that natural selection *is* an inventive force. Richard Dawkins, for example, assured his general readership that natural selection works as a "blind watchmaker" [4], and Graham Bell assured his technical readership that natural selection is, despite all the criticisms, one of "only two categories of theory competent to explain the living world," the other being design [5].

Adding further fuel to the controversy over natural selection is the undeniable fact that it does seem to accomplish some things quite well. Most significantly, perhaps, the natural production of high-affinity antibodies in the vertebrate immune system is achieved by a process that mirrors evolution on a small scale. Known as *affinity maturation*, the process uses repeated rounds of mutation (somatic hypermutation) and selection (B-cell clonal selection), organized within germinal centers of the lymph nodes [6]. Laboratory versions of molecular-selection have likewise proven capable of improving protein binding affinities [7] as well as thermal stabilities and catalytic activities [8]. Moreover, computer scientists and engineers have long recognized the utility of evolutionary strategies for solving certain optimization problems [9,10].

These two contrasting takes on selection can be assessed by recognizing that the evolution of biological inventions must involve two distinct steps. The first of these is the invention event itself—the initial coming together of the physical structures needed to produce a new adaptive function—and the second is the optimization of the invention by repeated mutation and selection of the genes that encode these structures. There is universal agreement that natural selection is not forward looking—it cannot bring things together in anticipation of the adaptive benefit that would result. This being so, the inability of selection to invent the things it optimizes is not controversial in itself.

The controversy seems instead to center on the significance of this limitation. Judging from human experience with technology, we naturally assign much greater difficulty to invention than we do to optimization. That is, experience tells us that the work of conceiving a new functional device, of refining this concept, and of implementing it in the form of a working prototype is considerably more challenging than the subsequent work of adjusting various aspects of the prototype to enhance performance or reduce costs. If this experience points to a universal truth—that the invention step is the hard step—then the inability of natural selection to invent does indeed leave a gaping hole in the middle of evolutionary theory, even after the optimizing power of selection is taken into consideration.

On the other hand, according to Darwin's notion of gradualism, every functional feature is thought to have been formed by processes that operated so gradually that there were no inventive discontinuities along the way. The evolution of the eye, for example, he imagined to have flowed in this seamless way [11]:

> If we must compare the eye to an optical instrument, we ought in imagination to take a thick layer of transparent tissue, with a nerve sensitive to light beneath, and then suppose every part of this layer to be continually changing slowly in density, so as to separate into layers of different densities and thicknesses, placed at different distances from each other, and with the surfaces of each layer slowly changing in form. Further we must suppose that there is a power always intently watching each slight accidental alteration in the transparent layers; and carefully selecting each alteration which, under varied circumstances, may in any way, or in any degree, tend to produce a distincter image. We must suppose each new state of the instrument to be multiplied by the million; and each to be preserved till a better be produced, and then the old ones to be destroyed. In living bodies, variation will cause the slight alterations, generation will multiply them almost infinitely, and natural selection will pick out with unerring skill each improvement.

The general idea here, which continues to influence evolutionary thinking, is that the many biological functions we see operating today all had rudimentary antecedents even before they were needed. If so, then the problem of invention disappears: the marvels of biology never had to be invented because they were always there in some crude form—they merely had to be optimized, and this is what natural selection does well.

Here, focusing in particular on protein functions, we consider two conceivable ways that these functions could have existed naturally before they were needed, so that they were ready for optimization the moment they were needed. The first possibility, which has become known as the *promiscuity hypothesis* [12], makes use of the observation that proteins that are well optimized for their primary functions may nevertheless have accidental side activities, much weaker than their primary ones. The thinking is that on rare occasions these minor activities prove beneficial, at which point a gene duplication event enables natural selection to optimize the newly beneficial function without loss of the original function. The common situation where an enzyme performs the same chemical transformation on multiple related substrates should not be confused with promiscuity, as the term is currently used. Mutations readily shift substrate preference in those common cases, but since the catalytic mechanism remains unchanged, shifts of that kind reveal nothing about enzyme origins.

Recognizing this, proponents of the promiscuity hypothesis have suggested that enzyme side activities should be described

as promiscuous only when two conditions are met. The first condition is that the side activity should have no physiological role, as otherwise the enzyme would be better described as *multifunctional* than as promiscuous [12]. The second condition has to do with the degree of separation between the side activity and the primary activity. Specifically, referring to the standard Enzyme Commission scheme for classifying enzyme functions[1], Khersonsky and Tawfik [12] recommend that the term "promiscuous" be reserved for cases where the EC numbers describing the primary and side activities differ in the first, second, or third index. For example, an enzyme acting primarily as a tryptophan *N*-monooxygenase (EC 1.14.13.125) would not be considered promiscuous if it exhibited the same activity to a lesser extent toward tyrosine (i.e., tyrosine *N*-monooxygenase activity; EC 1.14.13.41), but it would be considered promiscuous if it instead exhibited slight L-lysine 6-oxidase activity (EC 1.4.3.20) with no evident physiological role.

The second conceivable way for protein functions to have existed naturally before they were needed would be for "junk" proteins (i.e., proteins caused by accidental expression of non-genic DNA) to have weak non-specific activities, some of which might be optimized by natural selection when the need presents itself [13]. In this way even random open reading frames with no selective history might serve as starting points for the evolutionary optimization of pre-existing weak activities. The apparent disadvantage of the junk hypothesis relative to the promiscuity hypothesis is that junk proteins would have no stable folded structure at all. On the other hand, this might be construed as an advantage—in that the lack of commitment to any prior fold structure seemingly opens the door to invention of new structures.

The question to be addressed here is not whether promiscuous or junk activities exist in proteins or whether activities of these kinds can sometimes be selected. Rather, the question is whether the evolutionary optimization that occurs on those occasions when fortuitous activities are selected can explain the eventual appearance of highly proficient designs. That is, we are asking whether accidental activities can really be assumed to produce native-like activities by means of stepwise improvements. If we find this to be the case, then evolutionary invention is not as problematic as it seems. Conversely, if actual examples reveal no such connection between optimization and invention, this would argue that the problem of invention is real.

RESULTS

Selective optimization of weakly functional *Stylus* vector proteins

Stylus overview. We begin our investigation with a computational model called *Stylus* [14], which enables evolutionary experiments to be performed on artificial genes. *Stylus* does not aim to approximate the behavior of real genes or proteins but instead offers something analogous to these. At the core of the analogy is the rich set of real structure-to-function relationships associated with the written Chinese characters and their linguistic meanings. Analogous to the domain structure of proteins, Chinese characters may consist of multiple sub-structures in combination. And just as high-level functions of metabolic pathways and entire metabolomes require the coordinated actions of many proteins, so too high-level written communication requires the proper arrangement of many characters [14]. Reckoning a stand-alone text to occupy a position in this functional hierarchy roughly analogous to a complete proteome, we have used the free *Stylus* software [15] to generate the equivalent of a small genome, along with its encoded proteome [16].

The original *Stylus* paper [14] gives a complete description of how *Stylus* works, which we summarize here. *Stylus* uses twenty vectors as structural building blocks, analogous to the twenty amino acids. These vectors are joined end-to-end in the order specified by the open reading frame of a gene-like sequence. As in biology, the gene is translated one codon at a time with a genetic code specifying the vector encoded by each codon. Using *Stylus* terminology, we refer to these translated gene products as *vector proteins*. Once the vectors are joined to form a complete path, a simple drawing rule is applied to determine which portions of this path are drawn and which are undrawn. The ability to leave portions of the path undrawn enables a single *Stylus* gene to encode drawn structures with any number of visible strokes connected by invisible moves, a necessary condition for making Chinese characters.

Stylus calculates a numerical proficiency score (from 0 to 1) for a vector protein by comparing it to an ideal representation of a Chinese character, called an *archetype*. Human readers typically encounter illegibility when scores fall below about 0.1 (see figure 12 of reference 14), but scores are accurately handled down to much lower values, enabling experiments to be performed on vector proteins that have no noticeable resemblance to any character. It should be kept in mind, though, that while the various metrics used by *Stylus* to calculate scores were weighted according to their impact on human legibility[2], *Stylus* does not aim to model human character recognition. Rather, *Stylus* was simply inspired by the analogy between the Chinese characters and proteins, including the analogy between

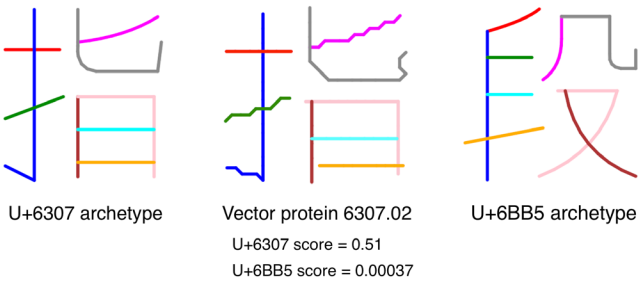

Figure 1: U+6BB5 (段) as a minor promiscuous function of vector protein 6307.02. Strokes (here individuated by color) are treated by *Stylus* as fundamental components of written characters [14], just as they are in actual writing. **doi:**10.5048/BIO-C.2015.2.f1

[1] See: http://www.chem.qmul.ac.uk/iubmb/enzyme/

[2] The original *Stylus* paper has a supplemental download (http://dx.doi.org/10.1371/journal.pone.0002246.s001) that describes the scoring algorithm in detail.

legibility and activity. This analogy informed the construction of the *Stylus* world with the intent that this artificial world would be interesting enough to merit study in itself.

Stylus experiments. The vector protein at the center of Figure 1 is a highly proficient version of 指 (Unicode index U+6307), which may mean *finger* or *to indicate*, depending on the context. The latter meaning applies in the context of the vector proteome [16] from which this vector protein (designated 6307.02) was taken. Although a human reader would not mistake 6307.02 for 段 (U+6BB5), which means *section*, there are nevertheless notable structural similarities between the two characters. First, as seen from the archetypes on either side of the vector protein in Figure 1, both characters are composed of three components arranged with a full-height component on the left and a pair of half-height components stacked on the right. In addition, both characters are drawn with nine strokes, as shown by line colors in the figure.

In the *Stylus* world, all that is needed for a vector protein to have measurable proficiency with respect to a particular character is a mapping of character strokes to drawn parts of the vector protein. So, by assigning the nine strokes in the 指-like vector protein 6307.02 to the nine strokes of 段 (U+6BB5), we can see how well this vector protein performs the 段 function. As shown for one plausible mapping (indicated by shared stroke colors between 6307.02 and U+6BB5 in Figure 1), 6307.02 performs this secondary function with a proficiency of 0.00037, which is more than three orders of magnitude lower than the 0.51 proficiency of its primary 指 function. This mirrors quite well the supposed starting point for evolutionary refinement of promiscuous activities. That is, while vector protein 6307.02 has strong 指 activity, it also has weak but quantifiable 段 activity, which, according to the promiscuity hypothesis, should enable this activity to be refined.

To test this hypothesis, we focused on simple Darwinian adaptive paths, meaning evolutionary paths where individual mutations are fixed sequentially by favorable selection. Although many more adaptive possibilities become theoretically possible if complex adaptations (where benefit requires two or more mutations) are considered, studies in population genetics have made it clear that fixation of complex adaptations is too rare in many higher organisms for this to provide a general route to new protein functions [17,18].

Stylus performs simple adaptive evolution by applying the steepest-ascent optimization algorithm as a built-in method option [15]. Using this algorithm, we tested whether the slight 段 activity of vector protein 6307.02 can, after optimization, approach the 0.51 proficiency standard met by all genes in the published *Stylus* genome [16]. Each round in the optimization started with a parent gene (the gene for 6307.02 being the first parent) from which all possible child genes that differ at one base position from their parent were produced and scored. The top-scoring child gene was then selected as the new parent. This process was automatically repeated until the parent was not outscored by any of its children, making this last parent gene locally optimal.

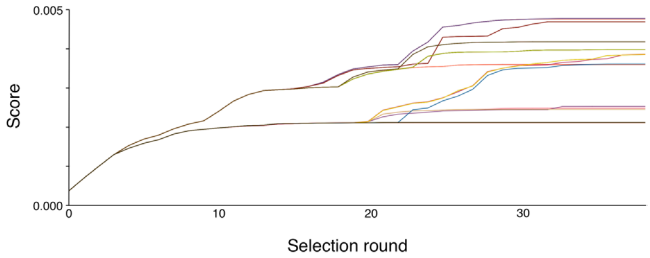

Figure 2: Evolutionary optimization of vector protein 6307.02 for the promiscuous U+6BB5 (段) function. Steepest-ascent optimization was performed forty times as described in Methods. Each curve therefore represents a plausible evolutionary trajectory, and while there are several of these, the fact that forty runs produced about nine distinguishable outcomes shows that the possibilities are very limited.
doi:10.5048/BIO-C.2015.2.f2

Figure 2 shows the results of forty independent steepest-ascent optimizations, many of which are overlapping. Although all the curves appear to plateau before reaching their final scores, an expanded vertical scale would show that slopes remain positive until optimization is complete. Branching is caused by the fact that multiple child genes (siblings) can have indistinguishably high scores[3]. Because *Stylus* chooses randomly among these indistinguishable siblings, the path of evolutionary optimization can take different routes on successive runs, even ending with different scores. None of these optimized scores, however, approach the 0.51 standard. The average score improvement is nearly tenfold, which may sound significant, but as seen in Figure 3, the initial score of 0.00037 (Figure 1) was so low that ten-fold improvement does not achieve anything comparable to the highly proficient examples of 段 from the published *Stylus* genome [16].

The next hypothesis to test with *Stylus* is the junk hypothesis. Again, the idea here is that a gene that does not specify any readable character at all might be a better starting point for

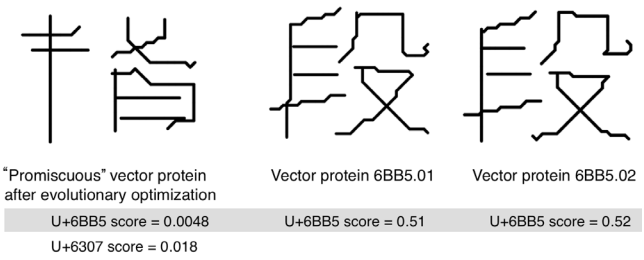

Figure 3: Ineffective evolutionary optimization of "promiscuous" U+6BB5 (段) function. The vector protein on the left is the best evolutionary outcome from the forty optimization trials shown in Figure 2. Two examples of vector proteins representing 段 from the published *Stylus* proteome [16] are shown for comparison. Notice that the optimization for 段 function has resulted in significant impairment of the original 指 function (U+6307), the new score for this function being about thirty-fold lower than the original score (see Figure 1).
doi:10.5048/BIO-C.2015.2.f3

[3] Selection is only able to fix the better of two competing genotypes if the selective advantage is greater than about $1/N_e$, where N_e is the effective population size [19,20]. If the difference is less than this, the two genotypes are indistinguishable with respect to fitness. The behavior of *Stylus* is consistent with these real-world facts.

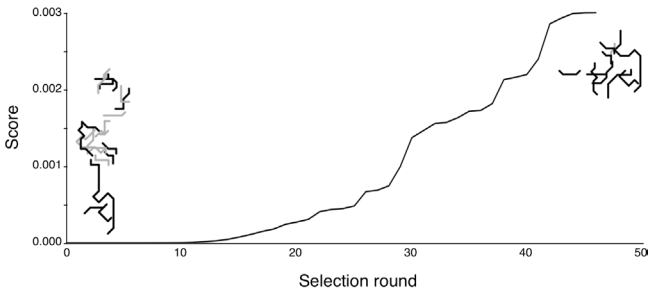

Figure 4: Ineffective evolutionary optimization of a junk vector protein for the U+6BB5 (段) function. From a starting score of 3×10^{-13}, steepest-ascent optimization reaches a final score of 0.003 at the 46th round, after which single base changes cause no further improvement. Grey lines in the vector proteins are extraneous marks consisting of drawn vectors that are not assigned to any of the nine strokes in 段. One of the noticeable effects of optimization is the near total elimination of marks, which elevates the score. **doi:**10.5048/BIO-C.2015.2.f4

evolutionary optimization of the 段 function than a gene that specifies some other function. For this test we obtained a junk gene by incorporating 3,000 successive random single-base substitutions into the 6BB5.02 gene (see Figure 3), using minimal selection to prevent loss of the open reading frame. At each step, all 1,251 possible base substitutions within the 417 base open reading frame were equally likely. Mutations were rejected only if they prevented scoring, in which case new mutations were generated to take their place. Because scoring is prevented only when a mutation causes a stroke to be lost, this minimal selective criterion guarantees the retention of nine strokes while allowing severe geometric deterioration.

The resulting vector protein (Figure 4, left) has an infinitesimal 段 score of $3.\times10^{-13}$ and, consistent with this, is wholly unrecognizable. This time, standard evolutionary optimization raises the 段 score to 0.003, which would seem very impressive if the starting point were the relevant standard. However, because we are asking whether proficient vector proteins comparable to those in the published vector proteome might be acquired by simple evolutionary optimization of junk vector proteins, the relevant standard is 0.51. As huge as the multiplicative improvement was, it stalled well below this 0.51 standard, which means evolutionary optimization failed to deliver a proficient vector protein. And while *Stylus* experiments are not based on visual recognition, the complete illegibility of the optimized vector protein (Figure 4, right) adds intuitive weight to this result.

Finally, to see whether a starting point that benefits from designed features would produce a more favorable outcome, we examined optimization of a vector protein that is well below the 0.51 proficiency standard but which has many of the structural aspects of a vector protein that meets this standard. We obtained this vector protein by repeatedly attempting to incorporate six random non-silent mutations into the 6BB5.02 gene until the resulting score was at least 0.00001. Although the result is a gene with a low score of 0.00029, the fact that only six out of the 139 vectors that form the vector protein have been replaced means that most of the structural features needed for high-level function are still intact. This can be seen in Figure 5 (lower left). In this case simple evolutionary optimization is strikingly successful, producing a vector protein with a 段 score of 0.54.

Selective optimization of weakly functional proteins

We performed two evolutionary optimization experiments in our laboratory that parallel the last two *Stylus* experiments just described. Before describing these, we briefly describe an experiment performed elsewhere that parallels the first of our *Stylus* experiments. Patrick and Matsumura [21] found that high-level overexpression of the *purF* gene in *Escherichia coli* compensates for the loss of tryptophan biosynthesis caused by deletion of the *trpF* gene, despite the completely different structures and chemistries of the two enzymes encoded by these genes. The *trpF* gene encodes phosphoribosyl anthranilate isomerase (PRAI; EC 5.3.1.24) which is classified by the SCOP database [22] as belonging to the TIM beta/alpha-barrel fold group[4], whereas *purF* encodes glutamine phosphoribosylpyrophosphate amidotransferase (EC 2.4.2.14), which has structural domains belonging to the Ntn hydrolase-like fold group[5] and the PRTase-like fold group[6]. The authors speculated that recognition of the phosphoribosyl moiety by both enzymes may be the basis of the apparent promiscuous activity (see Figure 6).

Of greater interest for the present study is the question of significance: Does the demonstration of very weak promiscuous PRAI activity in the product of *purF* imply that a proficient PRAI (comparable to the wild type) can be obtained by evolutionary optimization of this activity? After one round of evolutionary optimization in the laboratory, Patrick and Matsumura found that the best-performing *purF* mutant encoded

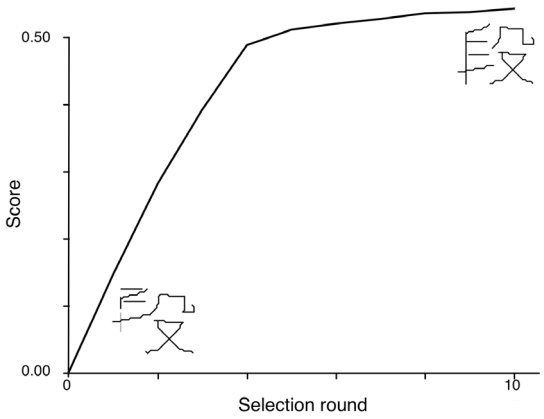

Figure 5: Highly effective evolutionary optimization of a sextuple mutant of 6BB5.02 for the U+6BB5 (段) function. From a starting score of 0.00029, steepest-ascent optimization reaches a final score of 0.54 after ten rounds of mutation and selection. The grey line in the vector protein on the left is an extraneous mark consisting of drawn vectors that are not assigned to any of the nine strokes in 段. **doi:**10.5048/BIO-C.2015.2.f5

[4] See: http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.b.c.e.html

[5] See: http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.e.dba.b.b.d.html

[6] See: http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.ih.b.b.h.html

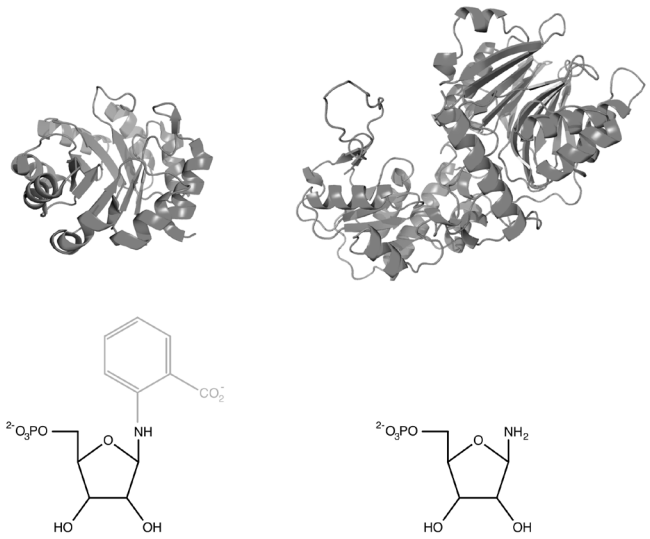

Figure 6: Proficient versus promiscuous catalysis of the PRAI reaction. PRAI, encoded by the *trpF* gene, is shown on left, with its substrate, *N*-(5-phosphoribosyl)-anthranilate, shown below. Glutamine phosphoribosylpyrophosphate amidotransferase, encoded by the *purF* gene, is shown on right, with its product, 5-phospho-β-D-ribosylamine, shown below. **doi:**10.5048/BIO-C.2015.2.f6

a protein with roughly 30-fold higher PRAI activity than the unmutated *purF* product. A second round of evolutionary optimization was unable to improve this result, implying that a local optimum had been reached. However, while this local optimum is measurably better than the non-mutant starting point, in vitro measurement of PRAI reaction kinetics showed the k_{cat}/K_M value for the evolved protein to be some *twenty-million-fold* lower than the value for the wild-type PRAI [21,23][7].

The numbers vary from case to case. Nevertheless, the inability of evolutionary optimization to reach wild-type catalytic proficiency from promiscuous starting points seems to be the rule rather than the exception. Schmidt and coworkers [24] reported a case of *near* promiscuity (promiscuous activity that appeared after a single nucleotide substitution) that was later found to have the same shortcoming. Their work began with the discovery that the L-Ala-D/L-Glu epimerase (AEE) from *E. coli* can catalyze the *o*-succinylbenzoate synthase (OSBS) reaction after one amino acid substitution (Asp297Gly) that requires a single base change [24]. Subsequent work by Vick and Gerlt showed that stepwise improvements of the promiscuous activity could be achieved by a succession of three more base changes [25]. However, the endpoint of these stepwise improvements had only 0.1% of wild-type OSBS activity [25], which while nearly 300-fold better than the promiscuous starting point, remains so low that the whole scenario seems to be of no actual evolutionary relevance. That is, the endpoint falls so far short of wild-type enzymes as to leave the origin of these enzymes unexplained.

Evolutionary optimization of a junk protein. Might a junk protein that happens to catalyze a chemical reaction provide a better starting point for evolutionary optimization than an enzyme with weak promiscuous function does? This proved not to be the case for the junk *Stylus* vector protein we examined, but perhaps real proteins behave differently in this regard. The thought that this may be so, anyway, has been invoked for evolutionary explanations [13] and has motivated searches for functional proteins among artificially generated libraries of random sequences [26].

To test whether invention from junk is any more likely for real proteins than it was for vector proteins, we performed laboratory evolutionary optimization on a previously described deletion protein [27] that enables bacteria to grow in the presence of low levels of ampicillin. Because this mutant was derived from the wild-type TEM-1 β-lactamase, we will refer to it as TEMΔ. As shown in Figure 7, TEMΔ is missing most of the original TEM-1 structure on one side of the active-site cleft. Whether the remaining polypeptide chain folds to form part of the wild-type TEM-1 structure is unknown, but the fact that much of the active site cleft (including important active-site residues Ser130 and Asn132) is missing implies that the ampicillin resistance cannot proceed by the normal TEM-1 mechanism. This was confirmed by finding that substituting Ser70 or Lys73 (both catalytically crucial residues [29]) has no

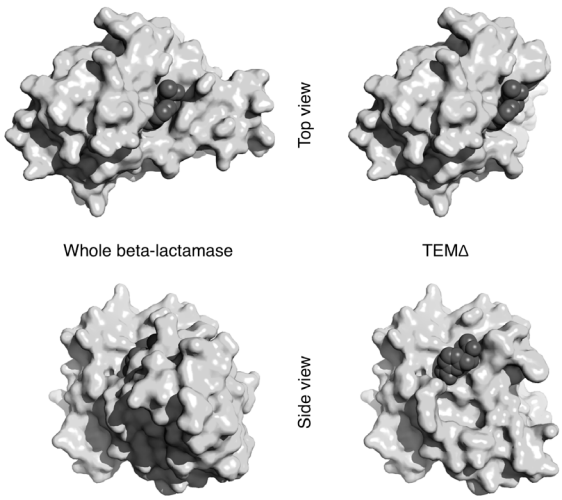

Figure 7: TEMΔ, a junk protein made by deleting 36 amino acid residues from the TEM-1 β-lactamase. Dark spheres show an inhibitor molecule bound in the active-site pocket where ampicillin normally binds. Structural representations of TEMΔ (right) are hypothetical in that residues from the wild-type structure (left) have simply been removed for the purpose of rendering. In addition to the 36 residues that are missing in the deletion mutant, another 29 residues are unequivocally prevented from adopting the wild-type conformation because of the missing segment, meaning that the structural disruption extends to 65 residues. Because this is the *minimum* extent of impact to the whole structure, the images on the right show the *maximum* amount of wild-type structure that could remain in TEMΔ. What actually remains may be much less. In addition to the deletion, TEMΔ carries 32 amino acid substitutions (see Supplement S1 [28]). **doi:**10.5048/BIO-C.2015.2.f7

[7] Patrick and Matsumura [21] say in their abstract that their evolved *purF* gene, despite its low activity, gives its possessor a "relative fitness" similar to that conferred by the PRAI gene. However, their Table 2 shows the growth rate of the evolved stain to be about one fifth that of the strain producing PRAI. In discussing this result, the authors do explain what they mean by relative fitness, but for our purpose of assessing the merits of the promiscuity hypothesis, the relevant fact is that their optimized *purF* gene product compares very poorly to PRAI.

Table 1: Ampicillin phenotypes observed during evolutionary optimization of TEMΔ

Round	0	1	2	3
Lowest concentration with no growth (μg/ml):	5–10	7	6–7	6
Highest concentration with growth (μg/ml):	5–7	5–7	6–7	5

effect on the low-level ampicillin resistance conferred by the TEMΔ [27].

Ampicillin resistance is typically reported in terms of the minimum inhibitory antibiotic concentration (MIC). The reported MIC for TEMΔ is 10 μg/ml at 25°C, which is about 5 μg/ml higher than the innate MIC for the commercial *E. coli* strain used (TOP10, Invitrogen) [27]. In the current work, we encountered batch-to-batch variation in the TOP10 strain that affected MIC measurements. Variability was more pronounced at higher MIC values (see below). Innate resistance MIC values for the various TOP10 batches ranged from 3–5 μg/ml, whereas MIC values for TEMΔ ranged from 5–10 μg/ml. In order to mitigate this effect, we re-measured important reference MIC values in parallel with experimental measurements using the same strain batch. This confirmed that TEMΔ does indeed consistently confer slightly higher than innate resistance, the increase being in the range of 2–5 μg/ml.

After three rounds of random mutagenesis and selection on agar growth medium with various ampicillin concentrations, we were unable to isolate a variant of TEMΔ with improved activity toward ampicillin (Table 1).

Evolutionary optimization of an impaired enzyme. The paper that describes the deletion mutant we are here referring to as TEMΔ [27] describes another derivative of the TEM-1 β-lactamase with 33 amino-acid substitutions but no deletions. Like TEMΔ, this mutant enhances bacterial survival in the presence of small amounts of ampicillin, but unlike TEMΔ this other mutant is sensitive to substitution of amino acids known to be essential for β-lactamase activity [27]. Since this property shows it to be a working β-lactamase enzyme, albeit a weak one, we will refer to this mutant as *basal-β*. The reported MIC for basal-β (previously referred to as the "reference sequence" [27]) is 20 μg/ml at 25°C, with the currently observed range being 15–45 μg/ml. Again, relative resistance remains consistent despite batch-to-batch variation in the host strain. That is, TEMΔ consistently causes slight enhancement of ampicillin resistance relative to the unassisted host strain, and basal-β consistently causes greater enhancement, though nowhere near the level of enhancement caused by the wild-type TEM-1 β-lactamase, which provides an MIC of 5,000 μg/ml at 25°C [27], or 6,600 μg/ml when tested with more recent batches of the TOP10 strain.

As shown in Table 2, this time successive rounds of random mutation and selection caused dramatic improvement, resulting in nearly complete recovery of the wild-type activity by the fourth and fifth rounds. In order to see whether it would be possible to exceed the performance of the wild-type enzyme, we increased the stringency of selection in the final round (round 6) by increasing the growth temperature from 25°C to 37°C, thereby requiring the enzyme to be stably folded at the optimal growth temperature of *E. coli*. The strongest mutant following this final selection, designated basal-β-mut6.4, has more than 500 times the activity of the basal-β starting point. More significantly, this optimized mutant is indeed more active against ampicillin than the wild-type TEM-1 enzyme, enabling growth above the wild-type MIC of 6,600 μg/ml at 37°C, and above 8,000 μg/ml at 25°C.

A sampling of amino acid substitutions present after each round of mutagenesis is shown in Figure 8. As seen in the final column of this figure and in the sequence alignment of Figure 9, basal-β_mut6.4 differs from the basal-β enzyme at 13 amino acid positions and from the wild-type TEM-1 enzyme at 37 amino acid positions. Of the 13 amino-acid changes accomplished by evolutionary optimization, three are reversions to the wild-type TEM-1 amino acid, three are non-wild-type substitutions at positions already carrying a non-wild-type amino acid in basal-β, and seven are further departures from the TEM-1 sequence.

DISCUSSION

An appreciation of what makes enzymes so remarkable will make the wider implications of this work more clear. As has been emphasized previously [30], while enzymes are proteins with catalytic activities, they are not *merely* this. The two catalytic proteins examined in this work, TEMΔ and basal-β, illustrate the distinction. Although the exact mechanism by which TEMΔ provides low-level ampicillin resistance is not known, it is reasonable to assume that this protein enhances catalytic hydrolysis of the antibiotic relative to the background rate. Ampicillin inhibits bacterial growth by interfering with biosynthesis of the peptidoglycan (or murein) layer of the cell wall [31]. Because this interference occurs outside the

Table 2: Ampicillin phenotypes observed during evolutionary optimization of basal-β

Round	0	1	2	3	4	5	6*
Lowest concentration with no growth (μg/ml):	10–45	n.d.	n.d.	3,000	n.d.	n.d.	n.d.
Highest concentration with growth (μg/ml):	5–40	70	800	2,000	6,000	6,000	8,000

* 8,000 μg/ml is the highest ampicillin concentration observed for basal-β-mut6.4 at 25°C, not the concentration at which this mutant was selected. As explained in the text, the final round of selection was performed at 37°C, whereas all prior rounds were performed at 25°C.

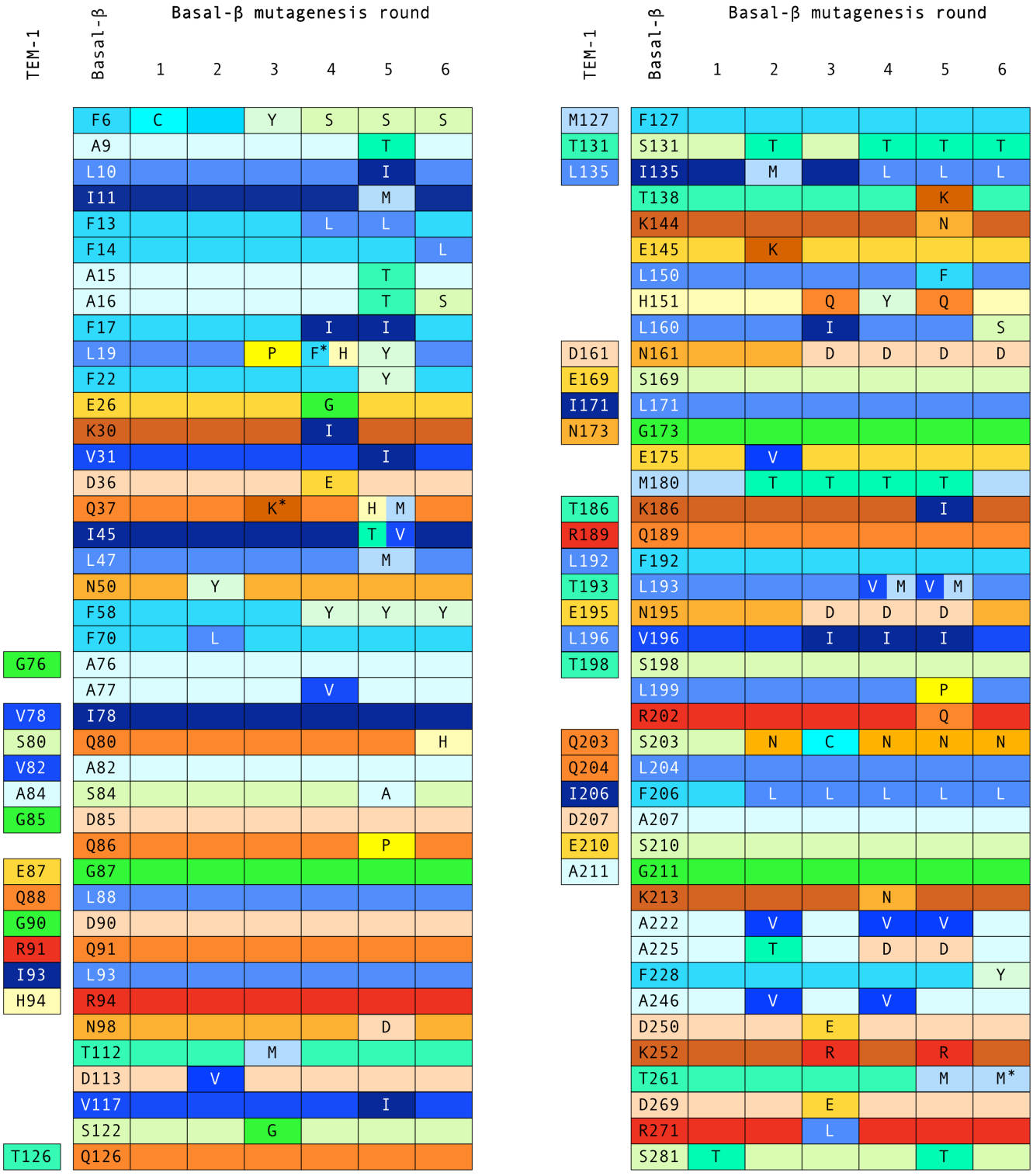

Figure 8: Amino acid substitutions found during evolutionary optimization of basal-β. Rows represent all amino acid positions with known substitutions, either in the initial basal-β sequence (in which case the wild-type TEM-1 amino acid is shown left) or after subsequent mutagenesis. Columns list amino acid substitutions present after each round of selection. Nine or ten clones were sequenced after all but the final round, where only the most active clone (basal-β-mut6.4) was sequenced. Except for round six, then, individual clones carried only a subset of the substitutions listed. Asterisks mark the three substitutions that have been reported among natural variants of the TEM-1 enzyme (see: http://www.uniprot.org/uniprot/P62593). The color scheme for representing the amino acids is taken from: http://www.imgt.org/IMGTScientificChart/RepresentationRules/colormenu.php. **doi:** 10.5048/BIO-C.2015.2.f8

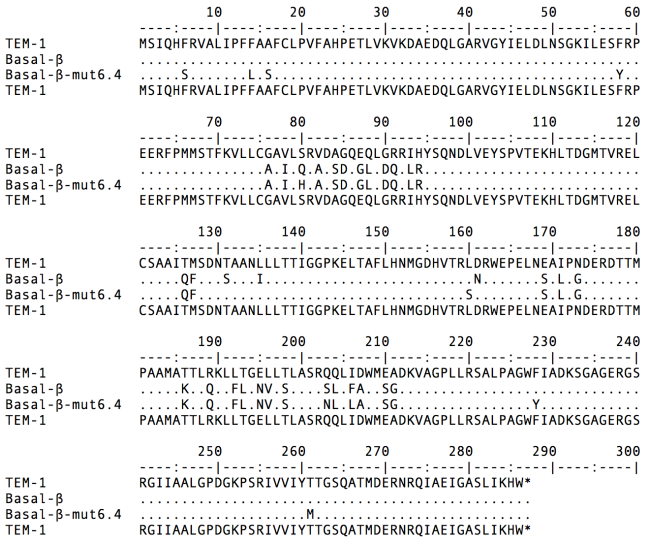

Figure 9: Alignment of basal-β and basal-β-mut6.4 amino-acid sequences to the wild-type TEM-1 sequence.
doi:10.5048/BIO-C.2015.2.f9

cytoplasmic compartment, natural β-lactamases have signal peptides that cause the mature enzymes to be exported from the cytoplasm [32]. The TEMΔ gene includes the natural upstream region coding for the TEM-1 signal peptide, making it likely that this mutant is exported in the usual way. As originally hypothesized [27], the susceptibility of ampicillin to hydrolysis by simple acid or base catalysis [33] suggests that TEMΔ may be doing nothing more than providing additional acidic and basic groups where they are most needed—in the immediate vicinity of the cell wall.

As true enzymes, β-lactamases hydrolyze ampicillin in a far more sophisticated way. Active sites, with their ability to exert precise chemical and geometric control over the local reaction environment, are what make this sophistication possible. The benefit is a dramatic increase in catalytic efficiency relative to simple catalysis, ranging from two to seventeen orders of magnitude for a wide range of unimolecular reactions [30,34].

In other words, it is the special arrangement of structural parts—and *only* this—that makes the exquisite performance of enzymes possible. Moreover, this seems to apply not just to enzymes but to everything else we recognize as having a specific function of any sophistication, whether from the living world or from the realm of human invention. What is true of enzymes is equally true of brains and eyes or of smartphones and sentences. We know that purposeful, intelligent action is necessary for achieving the special arrangement of letters that forms a coherent sentence or the special arrangement of materials that forms a smartphone, and the reasonableness of extending this knowledge to things like brains and eyes has been conceded even by people who believe that extension is ultimately mistaken. Geneticist Graham Bell, for example, acknowledged the point in these words [5]:

> A light bulb or a lathe are preconfigured in the mind, and constructed according to a plan. It is entirely reasonable to assume that beetles and daisies must be constructed after the same fashion, especially because they are much more complicated than anything that human ingenuity has so far managed to devise.

The title of Bell's book—*Selection: The Mechanism of Evolution*—leaves no doubt as to his favored alternative explanation of living things, but in placing the above words on his opening page, Bell also made it clear what selection must explain. The success of evolutionary theory requires not just that selection have real measurable effects but, more importantly, that the production of remarkable things—things beyond human ingenuity—be among those effects. In that regard, *the suggestion that all the remarkable things in the living world pre-existed is relevant only if the versions that might have pre-existed really could have been honed by selection into the impressive versions we now see.*

The computational *Stylus* experiments and the laboratory experiments described here provide a consistent picture of why this kind of evolutionary scenario fails. Both affirm the possibility of low-level functions being present even without structures that have been tailored to those functions. Furthermore, together with the study of Patrick and Matsumura [21], we have seen that these fortuitous low-level functions can be present either in structures that are tailored for *other* functions (the *purF* gene product having detectable PRAI function [21], and vector protein 6307.02 having detectable U+6BB5 function) or in junk sequences that may have no refined structure at all (TEMΔ providing some resistance to ampicillin and the vector protein of Figure 4 having quantifiable U+6BB5 function). This is not to say that all functions pre-exist in one of these fortuitous ways (indeed, it seems most do not[8]) but some clearly can.

Nevertheless, we consistently see that selection acting upon fortuitous functions of these kinds fails to produce the special structures—the *designs*—that would be needed to perform these functions well. The products of evolutionary optimization are proficient designs only if the necessary design elements were already present when selection was put into operation. This was apparent in the final *Stylus* optimization experiment, where the starting point was the sextuple mutant of 6BB5.02. Evolutionary optimization was able to exceed the 0.51 score standard in that case because, as shown in Figure 10, the essential structural elements were already there, needing only modest adjustment in order for their functional potential to be fully realized.

Our understanding of the connection between structure and function is far less complete for real proteins than it is for *Stylus* vector proteins, but the overall picture looks very similar. As previously reported [27], TEMΔ provides low-level ampicillin

[8] Patrick et al. [35] used 104 auxotrophic strains, each with a single gene knockout, and a plasmid library in which all *E. coli* genes are individually overexpressed to find out how many of the missing functions could be filled in by promiscuous rescue. Rescue was found to be possible for 21 of the knockouts, fifteen of them appearing to involve metabolic workarounds of various kinds and only six appearing to involve catalytic promiscuity [35], making the promiscuity rate under 6%. In the *Stylus* world, fortuitous representation of a character, whether by junk or by promiscuity, requires the source vector protein to have as many drawn parts as the character has strokes. As simple as this requirement is, it rules fortuitous function out in cases where it cannot be met.

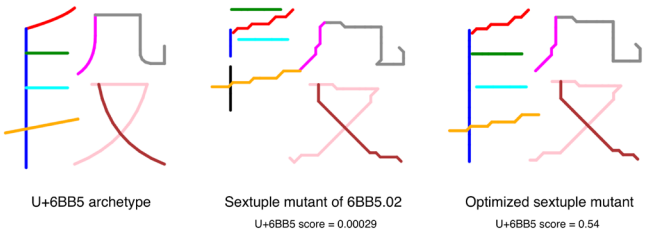

Figure 10: 6BB5.02 sextuple mutant retains most U+6BB5 structure. When strokes are individuated by color, we see that the sextuple mutant (center) has all nine strokes of the right shape with many in the right relative location and all but one (blue) being of the right size. *Stylus* treats the black vertical line in this mutant as an extraneous mark because it is not structurally contiguous with the blue stroke. This flaw, along with the resulting displacement of the green, cyan, and orange strokes, is successfully reversed by steepest-ascent optimization because everything is poised for substantial benefit to be realized one mutation at a time (Figure 5). **doi:**10.5048/BIO-C.2015.2.f10

resistance in a way that does not depend on the amino-acid side chains that are critical to the enzymatic function of the wild-type TEM-1 β-lactamase, whereas basal-β provides its low-level resistance in a way that does require these critical side chains. This distinction turns out to be much more significant than the phenotypic similarity of the strains that produce these two mutants. Since active-site side chains only take on their critical functions in the context of a complete active site, which only forms as part of a complete enzyme fold, the presence of functionally critical active-site residues indicates that the whole enzyme is present. In other words, the essential structural elements of the TEM-1 β-lactamase are present in basal-β but absent from TEMΔ, and this makes all the difference when it comes to evolutionary optimization. As with the sextuple mutant of 6BB5.02, only minor adjustments are needed in order for basal-β to display its full functional potential. TEMΔ, on the other hand, is like the junk vector protein of Figure 4, in that neither has any hidden potential.

We conclude that invention is not at all subsumed by optimization, contrary to Darwin's apparent thinking. If we define functions in terms of their measurable effects alone, paying no attention to the structures that produce them, then we are bound to encounter examples where some of these functions can be detected as slight effects in the absence of structures that are specific to them. As numerous studies have shown, it is often possible to devise circumstances (such as high-level gene expression) that enable these slight effects to be selected. However, considering the inherent limitations of natural selection in wild populations and the metabolic cost of amplifying slight effects by gene overexpression [36–38], these circumstances tend to be highly artificial—much more apt to be seen in a laboratory (where they have been arranged) than in a natural setting.

Moreover, we consistently find: 1) that pre-existing functions of this kind lack the underlying structural design that would be needed for them to be performed with high proficiency, and 2) that repeated rounds of mutation and selection are unable to invent that structure. If this structure has already been invented in some other way, then repeated mutation and selection can in some cases achieve large functional improvements by means of small structural adjustments. But this only explains how highly specific pre-existing structures can be fine-tuned, not how they originate. In other words, longstanding claims to the contrary notwithstanding, evolutionary theory seems to leaves biological invention wholly unexplained.

Interestingly, efforts in biotechnology to produce native-like enzymes with custom functions have led to similar conclusions. The hope has always been that evolutionary optimization can work wonders if supplied with even a crude design, but after decades of work it has become increasingly apparent that that initial design actually has to be remarkably good. A progress report from 2011 [39] concluded that:

> … efforts to date to generate novel catalysts have primarily demonstrated that we are getting good at making bad enzymes. Making good enzymes will require a whole new level of insight, or new methodologies altogether.

Indeed, good design always requires insight.

We have come to that conclusion not through studies in biotechnology but through studies in molecular biology and molecular evolution over the past fifteen years, beginning with a demonstration on two different enzymes that protein functions are much less tolerant of changes to their amino acid sequences than was commonly supposed [40]. That result raised the question of just how small the target within sequence space is that would need to be hit for a new enzyme fold to be produced. The next project [27] used limited randomization of the TEM-1 β-lactamase to answer this question, resulting in a deeper challenge to the Darwinian view. Estimated to amount to a mere 1 part in 10^{74} of the sequence possibilities, the target corresponding to a new functional enzyme fold is far too small to be hit by any known evolutionary process [27]. Building on this challenge, a full critique of the Darwinian explanation for protein folds was published in 2010 [41].

By that point we had turned our attention to evolutionary invention on a smaller scale. Having argued that the evolutionary process cannot come up with new protein folds, we asked whether it is able to invent new functions for existing folds. To explore this, we conducted an extensive search for amino acid substitutions that would enable any of a variety of enzymes within the GABA-aminotransferase-like family[9] to perform the function of one of their members—8-amino-7-oxononanoate synthase. Using both rational [42] and random [43] experimental approaches in conjunction with a mathematical analysis of complex adaptation in bacterial populations [44] we found evolutionary invention to be infeasible even on this small scale. Combining the fold-level results with results at the smaller scale, and considering all the relevant aspects of molecular evolution, we argued that the design interpretation of protein origins is much more plausible than the Darwinian interpretation [45].

Finally, we show here that invention is not superfluous. Although junk can be selected under some circumstances, the optimized endpoint after evolution has done all it can do is still junk. For the endpoint to be a highly refined functionally

[9] See: http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.d.jg.b.f.html

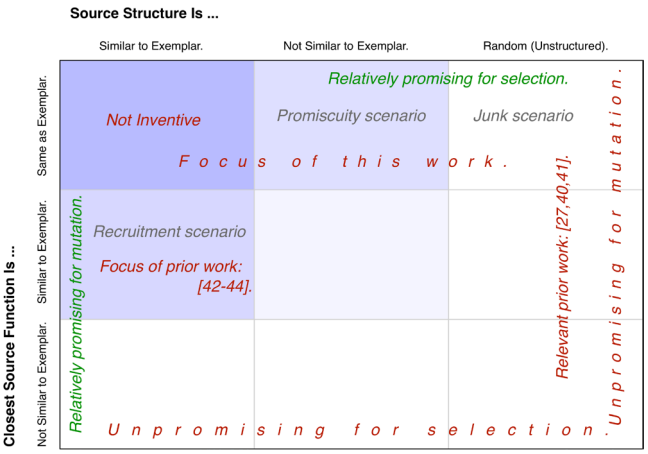

Figure 11: A grid for distinguishing protein-evolution scenarios. It is assumed here that the hypothesized source enzyme is not itself an exemplar for the function of interest. If it is, then there is no need to invent. Blue shading is strongest where the circumstances most favor significant evolutionary improvement and absent where circumstances seem to preclude this. Green and red text distinguish favorable considerations from unfavorable ones. **doi:**10.5048/BIO-C.2015.2.f11

specific structure, the essential aspects of that structure have to be present from the outset.

As an aid in seeing how these various projects form a coherent picture, Figure 11 places them on a grid representing the conceptual possibilities for the evolution of new enzymes. To make sense of this grid, remember that natural enzymes, with their extraordinary proficiencies, are what a successful theory of enzyme origins must explain. Given some natural enzymatic function of interest, we refer to a natural enzyme dedicated to this function as an *exemplar*. Now, according to the various evolutionary models, the starting point for the evolution of this exemplar was either a gene encoding some other well-structured protein (first two columns of Figure 11) or an open reading frame encoding an unstructured polypeptide (last column). In the first case, the source protein either did (column 1) or did not (column 2) have essentially the same fold structure as the exemplar. For each of these structural possibilities, then, we have a range of possible functional similarities between the source and the exemplar. By representing these possible combinations as distinct cases, the grid enables us to map the various evolutionary scenarios to different grid cells, as shown for the classical recruitment scenario, the promiscuity scenario, and the junk scenario.

Viewed in this way, we see that the challenge for evolutionary invention is that the one grid cell where evolution has everything it needs to work (heaviest blue shading) is the cell where neither the function nor the overall fold structure of the exemplar can be invented because both are already present. Conversely, in cells where invention is made conceptually possible by structural and/or functional differences from the exemplar, it seems to be made practically impossible by those very same differences. According to all the approaches we have used to examine the problem, there is no "sweet spot" where evolution does the work of invention. When pressed to work, evolution refines rather than invents, sometimes to significant effect, and when pressed to invent it does nothing of significance at all. Yet invention is precisely the work evolution must do—exquisitely well—if it is to explain life.

METHODS

Stylus experiments

Stylus experiments were performed with the freely available stand-alone *Stylus* application [15]. Each experiment involves running one or more *Stylus* methods on one or more *Stylus* genes. Here we provide more detailed description of some of the methods used, along with links to the actual *Stylus* files. As XML files, these are not intended to be intelligible to human readers, but by importing these to the *Stylus* application, users can examine them within the application and re-run the same experiments.

The gene for the first experiment was derived from gene 6307.02, which is available as part of the published genome [16]. The modified version of this gene file with strokes assigned to character U+6BB5 is available as a supplement to this paper (Supplement S2 [46]). The proficiency score of this gene was optimized by applying the steepest-ascent algorithm with "Factor" set to 1.00000001 (i.e., $1 + 10^{-8}$), corresponding to an effective population size of 10^8. The method file is available as Supplement S3 [47].

The second experiment used two *Stylus* methods. The first (Supplement S4 [48]) was used to radically degrade the 6BB5.02 gene from the published genome [16]. The method subjects the gene to 3,000 random base changes in succession, as described in Results. The resulting gene (Supplement S5 [49]) encodes the junk vector protein shown in Figure 4 (left). The method used for evolutionary optimization of this vector protein was the same as that used in the previous experiment (Supplement S3 [47]).

The degradation step of the final *Stylus* experiment achieved a low proficiency score by causing localized disruption at several locations. As described, this was done with a method that introduced six mutations at once (Supplement S6 [50]) and looking for a sextuple mutant with an appropriately low score. The chosen gene, which encodes the vector protein shown in Figure 5 (left), is available as Supplement S7 [51]. Optimization was performed as in previous experiments [47], with the resulting gene available as Supplement S8 [52].

Media and reagents

Cultures for plasmid preparation were grown in Terrific Broth with 20 μg/ml chloramphenicol (Sigma). The standard solid-culture medium was LB agar (Fluka) for strains without plasmids. All plasmid-bearing strains were maintained on LB agar with 20 μg/ml chloramphenicol (LBC20). All phenotype testing of plasmid libraries was done on LB plates with 7 μg/ml chloramphenicol (LBC7) and the specified concentration of ampicillin. Plates were freshly prepared the day before use to insure consistent ampicillin concentrations [27].

Mutazyme II kits and Pfu Ultra (Stratagene) were used for mutagenesis and PCR amplifications, respectively. PCR reagents were from NEB, and plasmid prep and gel purification kits were from Qiagen. Pure-and-Simple primers (Sigma) were used for PCR mutagenesis and amplification.

Plasmids and Strains

All experiments used *E. coli* strain TOP-10 (Invitrogen), which has the following published[10] genotype: F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G. Batch-to-batch variation in the ampicillin MIC values (discussed in the text) suggests uncharacterized genotypic variability at another locus, or possibly multiple loci. The plasmids carrying the wild-type TEM-1 gene and the mutant TEMΔ and basal-β genes are described in the original paper [27]. All plasmids carry the *cat* gene, which confers resistance to chloramphenicol.

Preparation and Screening of Mutagenized Libraries

Mutagenized gene libraries were prepared using Mutazyme II (Stratagene) with inwardly directed primers specific to the gene to be mutagenized. Conditions were adjusted to achieve tenfold amplification (estimated by gel analysis). After gel purification, the resulting mutant gene libraries were used as megaprimers for amplification of the entire expression vector, thereby replacing the unmutated TEMΔ or unmutated basal-β genes with the mutated libraries. Libraries were incorporated into the vector with high efficiency, with an average mutation rate of 1-2 mutations per kilobase (confirmed by sequencing). PCR products of the correct size were gel-purified, then circularized by ligation and used to transform TOP-10 cells by electroporation. Transformants were incubated for 2 hr at 30°C before diluting (1:10^2 to 1:10^4) and plating onto duplicate LBC7 plates (90 mm) for subsequent counting of viable transformants. The remaining cells were spread undiluted onto square LBC20 trays (245 mm; Becton Dickenson) for overnight growth at 37°C.

The next day, cells were washed from the tray in 5 mL LB, vortexed to create an even suspension, diluted 1:200 in LBC20, and incubated at 37°C for 6 hr at 250 rpm. Cells were then diluted 1:5000 for spreading onto medium with chloramphenicol and various amounts of ampicillin as described below. Higher dilutions (1:10^5 and 1:10^6) were again spread onto plates with chloramphenicol (no ampicillin) in order to establish how many cfu were plated.

[10] https://www.thermofisher.com/us/en/home/life-science/cloning/competent-cells-for-transformation/chemically-competent/top10f-genotypes.html

For each selection round, three pairs of plates were made, each pair containing LBC7 with one of three levels of ampicillin (see below). One half of each plate was spread with a 1:5000 dilution of cells from the current round (as described above), with the other half being spread with a 1:5000 dilution of cells revived from the previous round by thawing and growing overnight in LBC20 at 37°C, or a similarly cultured 1:5000 dilution of cells containing the wild-type TEM-1 enzyme. This enabled us to check the reproducibility of MIC measurements, and for the basal-β experiments it also enabled us to test each round at a larger range of ampicillin concentrations. The 1:5000 dilution typically yielded a density of 200,000-300,000 plasmid-carrying cells per half plate. After spreading, wrapped trays and plates were incubated for 42 hours at 25°C or 24 hours at 37°C and then scored for growth.

For each round of mutagenesis, the plate showing colonies at the highest ampicillin concentration was chosen as the starting point for the next round. Those colonies were washed from the plate and grown for six hours at 37°C shaking at 250 rpm. From these cultures three aliquots were frozen in 15% glycerol. Cells were also diluted 1:10^6, spread on LB C20, grown, and then streaked to get clonal colonies. These were used for sequencing of the β-lactamase gene (Seattle Biomedical Research Institute). Typically ten representatives from each round of mutagenesis were sequenced. The remainder of the 6-hour culture was plasmid-purified for use in the next round of mutagenesis and selection.

For each round of selection three levels of ampicillin were tested, one being the highest concentration that yielded growth in the previous round, the next being slightly higher, and the last being our best estimate (based on the vigor of the culture at the previous round) of the highest concentration likely to be within reach of cells carrying the newly mutagenized plasmids. In practice, ampicillin resistance increased so rapidly in the basal-β experiments that we rarely saw growth inhibition even at the highest ampicillin concentrations tested. TEMΔ, on the other hand, showed no improvement at all. We therefore chose the plate with colonies at the highest ampicillin concentration to try again, repeating nearly the same concentrations at each round.

Acknowledgements

The authors thank Winston Ewert for developing the *Stylus* stand-alone application used for the model portion of this work, and Stan Roberts for his capable assistance with the experimental portion.

1. Fontana W and Buss LW (1994) "The arrival of the fittest": Toward a theory of biological organization. B Math Biol 56:1–64.
2. De Vries H (1904) Species and Varieties– Their Origin by Mutation. Open Court Publishing (Chicago). De Vries attributes the quote to a Mr. Arthur Harris, without citing a source. I have added emphasis to underscore the point.
3. Wagner A (2014) Arrival of the Fittest—Solving Evolution's Greatest Puzzle. Current (New York). Emphasis in original, p 5.
4. Dawkins R (1986) The Blind Watchmaker. Longman (New York).
5. Bell G (1997) Selection—The Mechanism of Evolution. Chapman & Hall (New York). See Introduction.
6. Victoria GD, Nussenzweig MC (2012) Germinal centers. Ann Rev Immunol 30:429–457. **doi:** 10.1146/annurev-immunol-020711-075032

7. Barbas III CF, Wagner J (1995) Synthetic human antibodies: Selecting and evolving functional proteins. Methods 8:94–103. doi:10.1006/meth.1995.9997
8. Bloom JD, Arnold FH (2009) In the light of directed evolution: Pathways of adaptive protein evolution. Proc Natl Acad Sci USA 106:9995–10000. doi:10.1073/pnas.0901522106
9. Evolutionary Algorithms in Engineering Applications (1997) In: Dasgupta D, Michalewicz Z, eds. Springer-Verlag (Berlin).
10. Eiben AE, Smith JE (2003) Introduction to Evolutionary Computing. Springer-Verlag (Berlin, Heidelberg). doi:10.1007/978-3-662-05094-1
11. Darwin, CR (1859) The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, 1st ed. John Murray (London) pp 188–189.
12. Khersonsky O, Tawfik DS (2010) Enzyme promiscuity: A mechanistic and evolutionary perspective. Ann Rev Biochem 79:471–505. doi:10.1146/annurev-biochem-030409-143718
13. Begun DJ, Lindfors HA, Thompson ME, Holloway AK (2006) Recently evolved genes identified from *Drosophila yakuba* and *D. erecta* accessory gland expressed sequence tags. Genetics 172:1675–1681. doi:10.1534/genetics.105.050336
14. Axe DD, Dixon BW, Lu P (2008) Stylus: A system for evolutionary experimentation based on a protein/proteome model with non-arbitrary functional constraints. PLoS One 3(6): e2246. doi:10.1371/journal.pone.0002246
15. http://biologic.org/stylus
16. Axe DD, Lu P, Flatau S (2011) A Stylus-generated artificial genome with analogy to minimal bacterial genomes. BIO-Complexity 2011(3): 1–15. doi:10.5048/BIO-C.2011.3
17. Durrett R, Schmidt D (2008) Waiting for two mutations: With applications to regulatory sequence evolution and the limits of Darwinian evolution. Genetics 180:1501–1509. doi:10.1534/genetics.107.082610
18. Behe MJ, Snoke DW (2004) Simulating evolution by gene duplication of protein features that require multiple amino acid residues. Protein Sci 13: 2651–2664. doi:10.1110/ps.04802904
19. Kimura M (1955) Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harb Sym 20:33–53. doi:10.1101/SQB.1955.020.01.006
20. Kimura M (1994) Population Genetics, Molecular Evolution, and the Neutral Theory: Selected Papers. University of Chicago Press (Chicago) p 2.
21. Patrick WM, Matsumura I (2008) A study in molecular contingency: Glutamine phosphoribosylpyrophosphate amidotransferase is a promiscuous and evolvable phosphoribosylanthranilate isomerase. J Mol Biol 377:323–336. doi:10.1016/j.jmb.2008.01.043
22. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540. doi:10.1016/S0022-2836(05)80134-2
23. Hommel U, Eberhard M, Kirschner K (1995) Phosphoribosyl anthranilate isomerase catalyzes a reversible Amadori reaction. Biochemistry-US 34:5429–5439. doi:10.1021/bi00016a014
24. Schmidt DM, Mundorff EC, Dojka M, Bermudez E, Ness JE, et al (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. Biochemistry 42:8387–8393. doi:10.1021/bi034769a
25. Vick JE, Gerlt JA (2007) Evolutionary Potential of $(\beta/\alpha)_8$-barrels: Stepwise evolution of a "new" reaction in the enolase superfamily. Biochemistry-US 46:14589–14597. doi:10.1021/bi7019063
26. Keefe AD, Szostak JW (2001) Functional proteins from a random-sequence library. Nature 410:715–718. doi:10.1038/35070613
27. Axe DD (2004) Estimating the prevalence of protein sequences adopting functional enzyme folds. J Mol Biol 341: 1295–1315. doi:10.1016/j.jmb.2004.06.058
28. Supplement S1: TEM-1_TEMΔ_align.txt doi:10.5048/BIO-C.2015.2.S1
29. Strynadka NCJ, Adachi H, Jensen SE, Johns K, Sielecki A, Betzel C, et al (1992) Molecular structure of the acyl-enzyme intermediate in β-lactam hydrolysis at 1.7 Å resolution. Nature 359:700–705. doi:10.1038/359700a0
30. Axe DD, Foster NW, Fersht AR (1998) A search for single substitutions that eliminate enzymatic function in a bacterial ribonuclease. Biochemistry-US 37:7157–7166. doi:10.1021/bi9804028
31. Tomasz A (1979) The mechanism of the irreversible antimicrobial effects of penicillins: How the beta-lactam antibiotics kill and lyse bacteria. Ann Rev Microbiol 33:113–137. doi:10.1146/annurev.mi.33.100179.000553
32. Pugsley AP (1993) The complete general secretory pathway in Gram-negative bacteria. Microbiol Rev 57:50–108.
33. Raffanti EF, King JC (1974) Effect of pH on the stability of sodium ampicillin solutions. Am J Hosp Pharm 31:745–751.
34. Radzicka A, Wolfenden R (1995) A proficient enzyme. Science 267:90–93. doi:10.1126/science.7809611
35. Patrick WM, Quandt EM, Swartzlander DB, Matsumura I (2007) Multicopy suppression underpins metabolic evolvability. Mol Biol Evol 24:2716–2722. doi:10.1093/molbev/msm204
36. Wagner A (2005) Energy constraints on the evolution of gene expression. Mol Biol Evol 22: 1365–1374. doi:10.1093/molbev/msi126
37. Stoebel DM, Dean AM, Dykhuizen DF (2008) The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. Genetics 178:1653–1660. doi:10.1534/genetics.107.085399
38. Gauger AK, Ebnet S, Fahey PF, Seelke R (2010) Reductive evolution can prevent populations from taking simple adaptive paths to high fitness. BIO-Complexity 2010(2):1–9. doi:10.5048/BIO-C.2010.2
39. Brustad EM, Arnold FH (2011) Optimizing non-natural protein function with directed evolution. Curr Opin Chem Biol 15: 201–210. doi:10.1016/j.cbpa.2010.11.020
40. Axe DD (2000) Extreme functional sensitivity to conservative amino acid changes on enzyme exteriors. J Mol Biol 301: 585–595. doi:10.1006/jmbi.2000.3997
41. Axe DD (2010) The case against a Darwinian origin of protein folds. BIO-Complexity 2010(1):1–12. doi:10.5048/BIO-C.2010.1
42. Gauger AK, Axe DD (2011) The evolutionary accessibility of new enzyme functions: A case study from the biotin pathway. BIO–Complexity 2011(1): 1–17. doi:10.5048/ BIO-C.2011.1
43. Reeves MA, Gauger AK, Axe DD (2014) Enzyme families—Shared evolutionary history or shared design? A study of the GABA-aminotransferase family. BIO-Complexity 2014(4):1–16. doi:10.5048/BIO-C.2014.4
44. Axe DD (2010) The limits of complex adaptation: An analysis based on a simple model of structured bacterial populations. BIO-Complexity 2010(4):1–10. doi:10.5048/BIO-C.2010.4
45. Axe DD, Gauger AK (2013) Explaining metabolic innovation: Neo-Darwinism versus design. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC, eds. Biological Information: New Perspectives. World Scientific, pp 489–507.
46. Supplement S2: 6307-02-reassigned.gene doi:10.5048/BIO-C.2015.2.S2
47. Supplement S3: steepest-assent.xml doi:10.5048/BIO-C.2015.2.S3
48. Supplement S4: radical_degrade.xml doi:10.5048/BIO-C.2015.2.S4
49. Supplement S5: junk.gene doi:10.5048/BIO-C.2015.2.S5
50. Supplement S6: six_simultaneous.xml doi:10.5048/BIO-C.2015.2.S6
51. Supplement S7: sextuple_mutant.gene doi:10.5048/BIO-C.2015.2.S7
52. Supplement S8: optimized_sextuple.gene doi:10.5048/BIO-C.2015.2.S8