Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Structural Biology

Directed enzyme evolution: beyond the low-hanging fruit

Moshe Goldsmith and Dan S Tawfik

The field of directed evolution has progressed to the point where it is feasible to engineer enzymes for unnatural substrates and reactions with catalytic efficiencies and regio-specificity or stereo-specificity that rival those of natural enzymes. Here, we describe the conceptual and methodological advances that have enabled this progress. We address methodologies based on small libraries enriched with improved variants and carrying compensatory stabilizing mutations. Such libraries can be combined with low-throughput screens that provide high accuracy and directly target the desired substrate and reaction conditions, and thereby provide highly improved variants.

Address
Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, Israel

Corresponding author: Tawfik, Dan S (dan.tawfik@weizmann.ac.il)

Current Opinion in Structural Biology 2012, **22**:406–412

This review comes from a themed issue on **Engineering and design**

Edited by **Jane Clarke** and **William Schief**

For a complete overview see the Issue and the Editorial

Available online 12th May 2012

0959-440X/$ – see front matter, Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.sbi.2012.03.010

Introduction

Enzyme engineering, and directed enzyme evolution in particular, are rapidly advancing research fields that tackle applicative challenges as well as an intellectual one — do we understand enzymes well enough to make new enzymes that are as good as the natural ones? Here, we focus on recent methodological and conceptual advances that have extended the capabilities of enzyme engineers. These advances enable researchers to engineer a wider variety of enzymes and to bridge large gaps (orders-of-magnitude) in rate and selectivity between natural and evolved enzymes. This review aims to cover the most recent trends in directed enzyme evolution. Given the wealth of recent reviews that cover this field, we have highlighted our own views regarding the present advances and possible future developments (other recent reviews include [1–6], and reviews addressing related topics include [7–12]).

The first section of this review addresses the notion that directed enzyme evolution is ready to tackle high-hanging fruit. The second section highlights conceptual advances derived from fundamental studies of protein evolution and their contribution to protein engineering. The final two sections address specific advancements in library making, screening, and selection methodologies.

Beyond the low-hanging fruit

Initial efforts in directed evolution enabled modest improvements in only certain enzymatic properties. As in any other emerging field, this usually meant low-hanging fruit in terms of the target enzyme, target substrate, and rate improvements. Many early advances and newly developed approaches remained as proof of principle demonstrations, and were not successfully applied to other proteins. This was particularly apparent with complex screening or selection systems that were technically challenging. Nonetheless, eventually, advances in library making and screening techniques, the combination of directed evolution with computational design, and foremost, a better understanding of the mechanisms of natural protein evolution, enabled researchers to push the 'engineering ceiling' — that is, to address substrates and reactions for which no natural enzymes are known, to obtain large improvements in the evolved catalytic activity (relative to the starting point, and *in* absolute k_{cat}/K_M values), and to alter specificity, stability, and other enzymatic properties.

Specifically, the catalytic efficiency of poorly active starting points has been improved by as much as 10^5-fold, thereby reaching k_{cat}/K_M values that approach the level of natural enzymes ($\geq 10^5$ M^{-1} s^{-1}; for the distribution of kinetic parameters of natural enzymes see [13]). Examples include the increase of several hundred folds in the activity of a T7 RNA polymerase with new promoters and nucleotide specificities by 200 rounds of directed evolution in a continuous system [14]. The $\sim 10^5$-fold increase in the catalytic efficiency of serum paraoxonase 1 (PON1), which was optimized toward the hydrolysis of an organophosphate by 8 rounds of mutation and selection, from a k_{cat}/K_M value of <3.3 M^{-1} s^{-1} in wild-type to $>2.9 \times 10^5$ M^{-1} s^{-1} in the evolved mutant [15]. The improvement by >2000-fold of a computationally designed Kemp Eliminase, to a k_{cat}/K_M value of 0.6×10^6 M^{-1} s^{-1} following 16 rounds (Khersonsky O *et al.*, unpublished data). And finally, the long-term endeavor of converting a naturally occurring fatty acid P450 hydroxylase into a laboratory-evolved propane mono-oxygenase with a k_{cat}/K_M of 4.4×10^4 M^{-1} s^{-1} (only fourfold lower than wild-type with laurate). The wild-type enzyme does not hydroxylate propane at all, and the

Petitioner Merck, Ex. 1141, p. 406

first intermediate along this evolutionary trajectory that showed detectable activity with propane had to be improved by a factor of ca. 10^4 [16].

Dramatic improvements in the stereo-specificity and regio-specificity of evolved enzymes have also been reported. The engineer's challenge in such cases is to shift selectivity not only by reducing the catalytic efficiency of the evolved enzyme with the undesired, or the original substrate (as is often reported), but primarily to improve efficiency with the target substrate. Such improvements are therefore manifested in k_{cat}/K_M value for the target substrate and in the ratio of k_{cat}/K_M values for the target substrate over the undesired, or the original substrate. The k_{cat}/K_M ratio is directly proportional to the ratio of conversion rates at all substrate concentrations, and therefore comprises the standard for comparison. Recent examples of evolved enzymes with significant inversions in stereo-specificity include: a limonene epoxide hydrolase (LEH) [17], a transaminase [18], a lipase [19], a haloalkane dehalogenase [20], and serum PON1. In the latter case, for example, the barely detectable activity with the S isomer was improved by $\sim 3 \times 10^3$-fold to a k_{cat}/K_M of 0.85×10^6 M^{-1} s^{-1}. In doing so, wild-type's k_{cat}/K_M ratio for the two stereoisomers ($E_{(S/R)} < 6.3 \times 10^{-4}$) was completely inverted to a ratio of $E_{(S/R)} > 2500$ (Goldsmith M *et al.*, unpublished data).[a]

Enzymes can be evolved to catalyze reactions and/or substrates with which the parental wild-type protein shows no detectable activity. Examples include the hydroxylation of propane by P450 [16] or of butane by an evolved monooxygenase [21], and the hydrolysis of the nerve agent VX by an evolved PON1 variant (Goldsmith M *et al.*, unpublished data). Dramatic changes in thermostability [22–24], compatibility with organic solvents [18,25], or resistance to peroxide and low pH [26] have also been obtained.

Another challenge facing protein engineers regards *de novo* enzymes. The objective can be either to engineer an enzyme for a novel reaction (a reaction that is not catalyzed by natural enzymes) or to create a novel active-site, that is, by engineering an active-site from a noncatalytic site, or by fundamentally changing the active-site chemistry. In fact, evolving a completely new chemistry is an endeavor that is also rarely taken by nature. This is evident from the fact that the catalytic chemistry is conserved within enzyme-superfamilies while substrate specificity (and to a lesser degree reaction specificity) varies [27]. Consequently, catalytic promiscuity is much more common than substrate ambiguity [28]. The scope of *de novo* engineering has considerably expanded through advents in computational protein design. However, the catalytic efficiencies of *de novo* designed enzymes are still very low and they mostly catalyze highly activated substrates and simple reactions [29] (for recent reviews see [3,9]). Nonetheless, it appears that designed enzymes are highly evolvable, and can be refined by directed evolution in the same manner applicable to promiscuous activities of natural enzymes [30–33]. So although at its infancy, computational design, combined with directed evolution, holds significant promise in the area of *de novo* engineering.

Protein engineers turned evolutionists

Our own view is that advancements in enzyme engineering are rooted not so much in technological breakthroughs but more so in conceptual advancements, and in particular in a better understanding of how proteins evolve [34]. DNA shuffling is possibly the first example of borrowing from nature the highly successful evolutionary strategy of homologous recombination [35]. More recent examples include the recognition that new functions evolve by augmenting weak, promiscuous functions in existing proteins, and that the potential for evolving new functions can be enhanced by neutral mutations [34,36,37]. Implementation of the latter to laboratory evolution involves random mutagenesis and purifying selection for the enzyme's original, native activity. As neutral mutations are >100-fold more frequent than adaptive ones [38], medium-throughput screens can yield a diverse ensemble of variants that have accumulated neutral mutations. These 'neutral drift' ensembles can be subsequently screened for the new enzymatic activity [39,40], or for enzymes with reduced antigenicity [41].

The understanding that protein stability is a major limiting factor in the acquisition of new functions has also been crucial [42,43]. Most mutations that alter enzymatic activity are also destabilizing. Such mutations may not be tolerated unless accompanied by stabilizing, compensatory mutations (see ancestral–consensus mutations below) or buffered by chaperones [44]. This trend is especially noticeable in the application of many rounds, whereby the evolving protein's stability becomes severely compromised [16] (Goldsmith M *et al.*, unpublished data). It is also applicable to the directed evolution of human enzymes that tend to be marginally stable [45], and even to computationally designed enzymes (Khersonsky O *et al.*, unpublished data).

Other insights from natural evolution regard mutational pathways. The effects of mutations are often nonadditive, or epistatic. In particular, many mutations can be neutral or even deleterious but become advantageous in combination with others mutation. Non-additivity greatly complicates the search for the optimal set of mutations that confers the highest activity in a given enzyme. Although many trajectories are gradual and therefore unravel via single mutational steps, better understanding of epistatic effects is of crucial importance for successful engineering [23,42,46].

[a] Goldsmith M, Ashani Y, Simo Y, Ben-David M, Leader H, Silman I, Sussman JL, Tawfik DS: **Evolved stereoselective hydroalses for broad-spectrum G-type nerve agent detoxification**. *Chem Biol* 2012, in press.

Petitioner Merck, Ex. 1141, p. 407

Library methodologies

Smaller and smarter

Our understanding and prediction power of protein structure–function relationships are still rudimentary. As a result, the generation and examination of a library of randomly mutated variants is the method of choice for many laboratory evolution experiments aimed at enzyme engineering. However, this approach usually demands high-throughput screens that are not applicable in many cases. Because the frequency of beneficial mutations in such random libraries is very low (in the order of 10^{-3}), and >30% of mutations have deleterious effects, the accumulation of random mutations rapidly renders a protein completely inactive [38]. To increase the frequency of folded, active library variants, and particularly of variants with new functions, researchers are generating libraries that targeted specific regions of proteins and sequence space and filter or compensate for deleterious mutations. To this end, structural and evolutionary data are integrated, and sometimes computation is applied, to give small yet smart libraries (see also 'neutral drift' libraries above). These libraries are characterized by a large fraction of active variants despite carrying a surprisingly large number of active-site mutations (e.g. see [15,18,47–50]). Such libraries can yield results that are comparable to several rounds of evolution with conventional random mutagenesis libraries [50].

Structure-guided mutagenesis

Rational library design assumes that enzymatic activities evolve through changes in active-site residues that mediate substrate binding and catalysis. It applies primarily when the desired activity already exists in the starting point enzyme as a promiscuous one, and when detailed structural information is available. In two recent examples, this approach was used to engineer a P450 mono-oxygenase, and a nucleotidyltransferase, by so-called 'semi-rationale design'. By examining all residues in the active-site within a given distance from the substrate, a small number of residues were targeted for saturation mutagenesis [51,52]. Targeted mutagenesis (of selected positions, and specific amino acids within these positions) was also applied toward the directed evolution of an organophosphate hydrolase [15] (Goldsmith M *et al.*, unpublished data).

Targeted mutagenesis can also be performed in an iterative, systematic manner [53,54]. On the basis of detailed structural information, this approach enabled the introduction of allostery and the increase of substrate specificity of a Bayer–Villiger mono-oxygenase [55]. However, although exploring one position at a time is highly economical in screening capacity, this approach may fail to access non-additive, epistatic combinations (see above). Parallel exploration of more than one position will identify epistatic combinations, but demands the screening of large libraries (saturated mutagenesis of five active-site positions (to all 20 amino acids) yields a library $>10^6$ variants). One way of overcoming this hurdle is by spiking the gene library with oligonucleotides encoding the randomized positions [56]. In this way, a large number of positions can be explored while individual library variants contain only few mutations (e.g. see [15,31,49,57,58]).

Computer algorithms can enable more informed decisions regarding the choice of positions and amino acids. Several algorithms for the identification of mutational hotspots and mutational combinations (correlated mutations) based on sequence, structure, and function information have been developed [59–62]. These have been used to guide the evolution of enantioselectivity [20,47] and thermostability [63]. In addition, statistical tools have been used to identify combinations of beneficial mutations that originally appeared in separate variants [64,65], or to simultaneously optimize for more than one trait [66].

Phylogenetic libraries

Phylogenetic analyses have become instrumental in guiding the construction of libraries enriched with active variants. Comparing the sequence of the target enzymes to sequences of related enzymes can guide library design [47]. A potentially more powerful approach is to examine the sequence of the ancestor from which different enzymes with different functions diverged. Typically, a phylogenetic tree is constructed that includes the target enzyme, its family members (orthologues), and families with related enzymatic activities (paralogues). The sequences of various nodes that connect paralogous families are predicted. The target enzyme is then diversified by combinatorially incorporating back-to-ancestor mutations at positions within or near the active-site region. Such ancestral libraries afforded a large variety of improved variants by screening <100 variants [48,49,67]. Ancestral mutations can also be included in standard rationally designed libraries (Goldsmith M *et al.*, unpublished data).

Certain ancestral mutations also act as stabilizing mutations that facilitate the acquisition of new functions by compensating for the destabilizing effects of functional mutations. These often overlap with 'consensus mutations' that can be readily identified from sequence alignments. Restoring positions at which the target enzyme differs from most other family members to the consensus amino acid, usually increases the protein's stability, and thereby boosts its evolvability [68]. Stabilizing ancestral–consensus mutations have been applied to improve heterologous expression [69] and increase thermostability [70,71]. Foremost, ancestral–consensus mutations can be spiked into libraries to promote functional changes in evolving enzymes (Goldsmith M *et al.*, unpublished data), including cases where no improvements could be observed in libraries before the addition of such stabilizing mutations (Khersonsky O *et al.*, unpublished data).

Petitioner Merck, Ex. 1141, p. 408

Beyond point mutations

In nature, genetic diversification goes far beyond point mutations and includes insertions and deletions, homologous and nonhomologous recombination, and other global gene rearrangements. There have been several recent attempts to include these in the arsenal of enzyme engineering methodologies. The recombination of proteins with low sequence identity can be performed using predetermined crossover points chosen by SCHEMA [72] and other algorithms [73–75] that minimize the frequency of nonfunctional variants. Other diversification methods include split proteins [76] and circular permutation whereby the protein's termini are connected and new ends are randomly generated [77].

Insertions and deletions, particularly in active-site loops, are the key to the emergence of new enzyme functions in nature. However, thus far, their successful implementation to directed enzyme evolution has been limited. Progress in this area includes active-site loop libraries [78] or combinatorial fragment exchange [79], and insertions of new domains to an existing enzyme scaffold (as well as shorter insertions and deletions) using a transposon [80]. Insertions by duplications are the most common type of naturally incorporated insertions. To imitate this diversification mode, a method for generating randomly located insertions by short tandem repeats (one to several amino acids) has been developed (Kipnis Y *et al.*, unpublished data).[b]

Despite the above-described progress, backbone remodeling, and the remodeling of active-site loops in particular, are rarely successful. This is true for library-based approaches and for rational and computational design. Relative to point mutations, backbone changes can induce dramatic changes in enzyme function, but at the same time, are also more likely to be deleterious. Better knowledge of how backbone changes occur in nature, and computational tools for backbone redesign (e.g. [81–84]) may therefore prove particularly useful.

Screening and selection methodologies

To screen or select?

The main dilemma facing those designing directed enzyme evolution experiments is whether to apply a high, or even ultrahigh throughput selection methodology (10^6 to $>10^{12}$ variants processed in parallel), or a screen that might be of much lower throughput (10^2–10^4 variants tested individually). Ultra or high throughput screens or selections become advantageous with libraries of variants generated by extensive random mutagenesis or by simultaneous saturation mutagenesis of many positions. Such libraries need to be applied when: firstly, the available phylogenetic and/or structural information is insufficient for the design of small and focused libraries. Secondly, the identity and position of the mutations required to induce the desired property are hard to predict. Thirdly, the target change in activity depends on highly epistatic mutations that demand the simultaneous exploration of more than one position. In these, and possibly in other cases, large libraries might be the method of choice.

Recent advances in ultrahigh throughput selection technologies and their applications include selections *in vivo* for enzymes that degrade propane or butane [21] or cellobiose [85], and a continuous system for the evolution of T7 RNA polymerase [14] (see also [10]). Degradation tags that enforce a low enzyme concentration and therefore select for higher specific activities can greatly increase the dynamic range of *in vivo* selections [86]. Examples of directed evolution by *in vitro* selections include mRNA-display [87] and compartmentalization in emulsion droplets (IVC) [88–90].

The merits of screening

As powerful as they may be, selection methodologies have inherent disadvantages. They need to be tailored for a given enzyme and activity, and setting them up is usually a project in itself. Unlike screens, where the activity of each and every library variant becomes known, selections only yield (or not) improved variants. They are therefore harder to control and more prone to artifacts. Foremost, selections, and particularly *in vitro* selections, usually make use of surrogate substrates (e.g. fluorogenic analogues) and/or modified reaction conditions (e.g. single turnover).

Most screens have low to medium-throughput (10^2–10^4 clones/round), but can utilize almost any biochemical or biophysical detection method including HPLC and NMR. Despite their lack of elegance, screens often 'deliver the goods' and may ultimately save time and labor (see 'the Okazaki maneuver' in [91]). Screens also have advantages: firstly, rapid assessment of the diversification strategy — for example, if >90% of variants show no activity with the enzyme's original substrate — change strategy. Secondly, the ability to monitor multiple parameters, for example, monitoring the original as well as the evolved activity. This enables the isolation of variants with higher expression levels versus variants with improved enzymatic activity, target higher activity as well as selectivity, and so on. Thirdly, precision of activity measurements, thus enabling the identification of marginally improved variants, or even neutral ones, that can later yield higher improvements in combination with additional mutations. Fourthly, nearly unlimited dynamic range — the screened variants only need to be further diluted as they become more improved. Fifthly, the actual substrate and desired reaction conditions can be applied, in short, 'you get what you select for' [42].

[b] Kipnis Y, Dellus-Gur E, Tawfik DS: **TRINS: a method for gene modification by randomized tandem repeat insertions**. *Protein Engineering, Design & Selection: PEDS* 2012, in press.

Petitioner Merck, Ex. 1141, p. 409

Screening for what one aims for is in our view a crucial factor and a significant element in 'high fruit' challenges. For example, when evolving enzymes for nerve agent detoxification, we ultimately had to move from ultrahigh throughput screens with fluorogenic surrogates to a medium-throughput assay. The screened variants and acetylcholine esterase (which the agents aim to block) were exposed to physiological nerve agent concentrations (<1 μM), and variants capable of intercepting the agent were detected by measuring the residual levels noninhibited acetylcholine esterase. This assay is rather complicated, and also demanded the development of suitable substrates (*in situ* synthesis of nerve agents at very low, nonhazardous concentrations). However, it isolated variants that hydrolyze the toxic component of these agents only (the S_P isomer of from a racemic mixture), at physiologically relevant concentrations, and fast enough to provide prophylactic protection [15] (Goldsmith M *et al.*, unpublished data).

Screens do not necessarily impose low-throughput. Thus, in principle, the merits of high-throughput and screening accuracy can be combined as indicated by the use of FACS (fluorescence activated cell sorter) for screening enzyme libraries using cell entrapped substrates [92,93], by compartmentalization in emulsion droplets [15,94,95], or by cell-surface display [96–98]. Albeit, FCAS screens usually demand the use of surrogate substrates and/or specialized expression systems. Other technological advents such as microfluidics may also afford high throughput while maintaining high sample control, accuracy, and the ability to monitor in parallel multiple parameters [99–102].

Acknowledgements

Financial support by the Defense Threat Reduction Agency (DTRA) of the US Department of Defense and by the CounterACT Program of the US National Institutes Of Health, by the Israel Science Foundation, by the Willner Albert and Blanche and the Jack Wolgin Foundations, are gratefully acknowledged.

References

1. Wang M, Si T, Zhao H: **Biocatalyst development by directed evolution**. *Bioresour Technol* 2012. (Epub ahead of print).
2. Dalby PA: **Strategy and success for the directed evolution of enzymes**. *Curr Opin Struct Biol* 2011, **21**:473-480.
3. Golynskiy MV, Seelig B: **De novo enzymes: from computational design to mRNA display**. *Trends Biotechnol* 2010, **28**:340-345.
4. Labrou NE: **Random mutagenesis methods for in vitro directed enzyme evolution**. *Curr Protein Pept Sci* 2010, **11**:91-100.
5. Lutz S: **Beyond directed evolution — semi-rational protein engineering and design**. *Curr Opin Biotechnol* 2010, **21**:734-743.
6. Jackel C, Hilvert D: **Biocatalysts by evolution**. *Curr Opin Biotechnol* 2010, **21**:753-759.
7. Bonsor DA, Sundberg EJ: **Dissecting protein–protein interactions using directed evolution**. *Biochemistry* 2011, **50**:2394-2402.
8. Baird CL, Fischer CJ, Pefaur NB, Miller KD, Kagan J, Srivastava S, Rodland KD: **Developing recombinant antibodies for biomarker detection**. *Cancer Biomark* 2010, **6**:271-279.
9. Brustad EM, Arnold FH: **Optimizing non-natural protein function with directed evolution**. *Curr Opin Chem Biol* 2011, **15**:201-210.
10. Bottcher D, Bornscheuer UT: **Protein engineering of microbial enzymes**. *Curr Opin Microbiol* 2010, **13**:274-282.
11. Behrens GA, Hummel A, Padhi SK, Schatzle S, Bornscheuer UT: **Discovery and protein engineering of biocatalysts for organic synthesis**. *Adv Synth Catal* 2011, **353**:2191-2215.
12. Portnoy VA, Bezdan D, Zengler K: **Adaptive laboratory evolution — harnessing the power of biology for metabolic engineering**. *Curr Opin Biotechnol* 2011, **22**:590-594.
13. 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.
14. Esvelt KM, Carlson JC, Liu DR: **A system for the continuous directed evolution of biomolecules**. *Nature* 2011, **472**:499-503.
15. Gupta RD, Goldsmith M, Ashani Y, Simo Y, Mullokandov G, Bar H, Ben-David M, Leader H, Margalit R, Silman I *et al.*: **Directed evolution of hydrolases for prevention of G-type nerve agent intoxication**. *Nat Chem Biol* 2011, **7**:120-125.
16. Fasan R, Meharenna YT, Snow CD, Poulos TL, Arnold FH: **Evolutionary history of a specialized P450 propane monooxygenase**. *J Mol Biol* 2008, **383**:1069-1080.
17. Zheng H, Reetz MT: **Manipulating the stereoselectivity of limonene epoxide hydrolase by directed evolution based on iterative saturation mutagenesis**. *J Am Chem Soc* 2010, **132**:15744-15751.
18. Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S, Jarvis WR, Colbeck JC, Krebber A, Fleitz FJ, Brands J *et al.*: **Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture**. *Science* 2010, **329**:305-309.
19. Sandstrom AG, Wikmark Y, Engstrom K, Nyhlen J, Backvall JE: **Combinatorial reshaping of the Candida antarctica lipase A substrate pocket for enantioselectivity using an extremely condensed library**. *Proc Natl Acad Sci U S A* 2012, **109**:78-83.
20. van Leeuwen JG, Wijma HJ, Floor RJ, van der Laan JM, Janssen DB: **Directed evolution strategies for enantiocomplementary haloalkane dehalogenases: from chemical waste to enantiopure building blocks**. *Chembiochem* 2012, **13**:137-148.
21. Koch DJ, Chen MM, van Beilen JB, Arnold FH: **In vivo evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6**. *Appl Environ Microbiol* 2009, **75**:337-344.
22. Zhang D, Zhu F, Fan W, Tao R, Yu H, Yang Y, Jiang W, Yang S: **Gradually accumulating beneficial mutations to improve the thermostability of N-carbamoyl-D-amino acid amidohydrolase by step-wise evolution**. *Appl Microbiol Biotechnol* 2011, **90**:1361-1371.
23. Gumulya Y, Reetz MT: **Enhancing the thermal robustness of an enzyme by directed evolution: least favorable starting points and inferior mutants can map superior evolutionary pathways**. *Chembiochem* 2011, **12**:2502-2510.
24. Mollwitz B, Brunk E, Schmitt S, Pojer F, Bannwarth M, Schiltz M, Rothlisberger U, Johnsson K: **Directed evolution of the suicide protein O6-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability**. *Biochemistry* 2012, **51**:986-994.
25. Reetz MT, Soni P, Fernandez L, Gumulya Y, Carballeira JD: **Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method**. *Chem Commun* 2010, **46**:8657-8658.
26. Garcia-Ruiz E, Gonzalez-Perez D, Ruiz-Duenas FJ, Martinez AT, Alcalde M: **Directed evolution of a temperature-, peroxide- and alkaline pH-tolerant versatile peroxidase**. *Biochem J* 2012, **441**:487-498.

Petitioner Merck, Ex. 1141, p. 410

27. Glasner ME, Gerlt JA, Babbitt PC: **Evolution of enzyme superfamilies**. *Curr Opin Chem Biol* 2006, **10**:492-497.

28. Khersonsky O, Malitsky S, Rogachev I, Tawfik DS: **Role of chemistry versus substrate binding in recruiting promiscuous enzyme functions**. *Biochemistry* 2011, **50**:2683-2690.

29. Baker D: **An exciting but challenging road ahead for computational enzyme design**. *Protein Sci* 2010, **19**:1817-1819.

30. Khersonsky O, Rothlisberger D, Dym O, Albeck S, Jackson CJ, Baker D, Tawfik DS: **Evolutionary optimization of computationally designed enzymes: Kemp eliminases of the KE07 series**. *J Mol Biol* 2010, **396**:1025-1042.

31. Khersonsky O, Rothlisberger D, Wollacott AM, Murphy P, Dym O, Albeck S, Kiss G, Houk KN, Baker D, Tawfik DS: **Optimization of the in-silico-designed kemp eliminase KE70 by computational design and directed evolution**. *J Mol Biol* 2011, **407**:391-412.

32. Wang L, Althoff EA, Bolduc J, Jiang L, Moody J, Lassila JK, Giger L, Hilvert D, Stoddard B, Baker D: **Structural analyses of covalent enzyme-substrate analog complexes reveal strengths and limitations of de novo enzyme design**. *J Mol Biol* 2012, **415**:615-625.

33. 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. (advance online publication).

34. Peisajovich SG, Tawfik DS: **Protein engineers turned evolutionists**. *Nat Methods* 2007, **4**:991-994.

35. Stemmer WP: **Rapid evolution of a protein in vitro by DNA shuffling**. *Nature* 1994, **370**:389-391.

36. Mannervik B, Runarsdottir A: **The quest for molecular quasi-species in ligand-activity space and its application to directed enzyme evolution**. *FEBS Lett* 2010, **584**:2565-2571.

37. Soskine M, Tawfik DS: **Mutational effects and the evolution of new protein functions**. *Nat Rev Genet* 2010, **11**:572-582.

38. 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.

39. Gupta RD, Tawfik DS: **Directed enzyme evolution via small and effective neutral drift libraries**. *Nat Methods* 2008, **5**:939-942.

40. 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.

41. Cantor JR, Yoo TH, Dixit A, Iverson BL, Forsthuber TG, Georgiou G: **Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift**. *Proc Natl Acad Sci U S A* 2011, **108**:1272-1277.

42. Romero PA, Arnold FH: **Exploring protein fitness landscapes by directed evolution**. *Nat Rev Mol Cell Biol* 2009, **10**:866-876.

43. Tokuriki N, Tawfik DS: **Stability effects of mutations and protein evolvability**. *Curr Opin Struct Biol* 2009, **19**:596-604.

44. Tokuriki N, Tawfik DS: **Chaperonin overexpression promotes genetic variation and enzyme evolution**. *Nature* 2009, **459**:668-673.

45. Amar D, Berger I, Amara N, Tafa G, Meijler MM, Aharoni A: **The transition of human estrogen sulfotransferase from generalist to specialist using directed enzyme evolution**. *J Mol Biol* 2011, **416**:21-32.

46. Reetz MT, Sanchis J: **Constructing and analyzing the fitness landscape of an experimental evolutionary process**. *Chembiochem* 2008, **9**:2260-2267.

47. Jochens H, Bornscheuer UT: **Natural diversity to guide focused directed evolution**. *Chembiochem* 2010, **11**:1861-1866.

48. Chen F, Gaucher EA, Leal NA, Hutter D, Havemann SA, Govindarajan S, Ortlund EA, Benner SA: **Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection**. *Proc Natl Acad Sci U S A* 2010, **107**:1948-1953.

49. Alcolombri U, Elias M, Tawfik DS: **Directed evolution of sulfotransferases and paraoxonases by ancestral libraries**. *J Mol Biol* 2011, **411**:837-853.

50. Chen MM, Snow CD, Vizcarra CL, Mayo SL, Arnold FH: **Comparison of random mutagenesis and semi-rational designed libraries for improved cytochrome P450 BM3-catalyzed hydroxylation of small alkanes**. *Protein Eng Des Sel* 2012, **25**:171-178.

51. Hoffmann G, Bonsch K, Greiner-Stoffele T, Ballschmiter M: **Changing the substrate specificity of P450cam towards diphenylmethane by semi-rational enzyme engineering**. *Protein Eng Des Sel* 2011, **24**:439-446.

52. Moretti R, Chang A, Peltier-Pain P, Bingman CA, Phillips GN Jr, Thorson JS: **Expanding the nucleotide and sugar 1-phosphate promiscuity of nucleotidyltransferase RmlA via directed evolution**. *J Biol Chem* 2011, **286**:13235-13243.

53. Reetz MT, Prasad S, Carballeira JD, Gumulya Y, Bocola M: **Iterative saturation mutagenesis accelerates laboratory evolution of enzyme stereoselectivity: rigorous comparison with traditional methods**. *J Am Chem Soc* 2010, **132**:9144-9152.

54. Reetz MT: **Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions**. *Angew Chem* 2011, **50**:138-174.

55. Wu S, Acevedo JP, Reetz MT: **Induced allostery in the directed evolution of an enantioselective Baeyer–Villiger monooxygenase**. *Proc Natl Acad Sci U S A* 2010, **107**:2775-2780.

56. Herman A, Tawfik DS: **Incorporating synthetic oligonucleotides via gene reassembly (ISOR): a versatile tool for generating targeted libraries**. *Protein Eng Des Sel* 2007, **20**:219-226.

57. Berger I, Guttman C, Amar D, Zarivach R, Aharoni A: **The molecular basis for the broad substrate specificity of human sulfotransferase 1A1**. *PLoS One* 2011, **6**.

58. Scanlon TC, Teneback CC, Gill A, Bement JL, Weiner JA, Lamppa JW, Leclair LW, Griswold KE: **Enhanced antimicrobial activity of engineered human lysozyme**. *ACS Chem Biol* 2010, **5**:809-818.

59. Treynor TP, Vizcarra CL, Nedelcu D, Mayo SL: **Computationally designed libraries of fluorescent proteins evaluated by preservation and diversity of function**. *Proc Natl Acad Sci U S A* 2007, **104**:48-53.

60. Pavelka A, Chovancova E, Damborsky J: **HotSpot wizard: a web server for identification of hot spots in protein engineering**. *Nucleic Acids Res* 2009, **37**:W376-W383.

61. Kuipers RK, Joosten HJ, van Berkel WJ, Leferink NG, Rooijen E, Ittmann E, van Zimmeren F, Jochens H, Bornscheuer U, Vriend G *et al.*: **3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities**. *Proteins* 2010, **78**:2101-2113.

62. Ma BG, Berezovsky IN: **The MBLOSUM: a server for deriving mutation targets and position-specific substitution rates**. *J Biomol Struct Dyn* 2010, **28**:415-419.

63. Cerdobbel A, De Winter K, Aerts D, Kuipers R, Joosten HJ, Soetaert W, Desmet T: **Increasing the thermostability of sucrose phosphorylase by a combination of sequence- and structure-based mutagenesis**. *Protein Eng Des Sel* 2011, **24**:829-834.

64. Brouk M, Nov Y, Fishman A: **Improving biocatalyst performance by integrating statistical methods into protein engineering**. *Appl Environ Microbiol* 2010, **76**:6397-6403.

65. Hokanson CA, Cappuccilli G, Odineca T, Bozic M, Behnke CA, Mendez M, Coleman WJ, Crea R: **Engineering highly thermostable xylanase variants using an enhanced combinatorial library method**. *Protein Eng Des Sel* 2011, **24**:597-605.

66. He L, Friedman AM, Bailey-Kellogg C: **A divide-and-conquer approach to determine the Pareto frontier for optimization of protein engineering experiments**. *Proteins* 2011, **80**:790-806.

Petitioner Merck, Ex. 1141, p. 411

67. Cole MF, Gaucher EA: **Exploiting models of molecular evolution to efficiently direct protein engineering**. *J Mol Evol* 2011, **72**:193-203.

68. Bershtein S, Goldin K, Tawfik DS: **Intense neutral drifts yield robust and evolvable consensus proteins**. *J Mol Biol* 2008, **379**:1029-1044.

69. Khersonsky O, Rosenblat M, Toker L, Yacobson S, Hugenmatter A, Silman I, Sussman JL, Aviram M, Tawfik DS: **Directed evolution of serum paraoxonase PON3 by family shuffling and ancestor/consensus mutagenesis, and its biochemical characterization**. *Biochemistry* 2009, **48**:6644-6654.

70. Zhang ZG, Yi ZL, Pei XQ, Wu ZL: **Improving the thermostability of Geobacillus stearothermophilus xylanase XT6 by directed evolution and site-directed mutagenesis**. *Bioresour Technol* 2010, **101**:9272-9278.

71. Jochens H, Aerts D, Bornscheuer UT: **Thermostabilization of an esterase by alignment-guided focussed directed evolution**. *Protein Eng Des Sel* 2010, **23**:903-909.

72. Meyer MM, Silberg JJ, Voigt CA, Endelman JB, Mayo SL, Wang ZG, Arnold FH: **Library analysis of SCHEMA-guided protein recombination**. *Protein Sci* 2003, **12**:1686-1693.

73. Zheng W, Friedman AM, Bailey-Kellogg C: **Algorithms for joint optimization of stability and diversity in planning combinatorial libraries of chimeric proteins**. *J Comput Biol* 2009, **16**:1151-1168.

74. Zheng W, Griswold KE, Bailey-Kellogg C: **Protein fragment swapping: a method for asymmetric, selective site-directed recombination**. *J Comput Biol* 2010, **17**:459-475.

75. Parker AS, Griswold KE, Bailey-Kellogg C: **Optimization of combinatorial mutagenesis**. *J Comput Biol* 2011, **18**:1743-1756.

76. Chen Y, Li S, Chen TJ, Hua H, Lin ZL: **Random dissection to select for protein split sites and its application in protein fragment complementation**. *Protein Sci* 2009, **18**:399-409.

77. Yu Y, Lutz S: **Circular permutation: a different way to engineer enzyme structure and function**. *Trends Biotechnol* 2011, **29**:18-25.

78. Ochoa-Leyva A, Barona-Gomez F, Saab-Rincon G, Verdel-Aranda K, Sanchez F, Soberon X: **Exploring the structure–function loop adaptability of a (beta/alpha)(8)-barrel enzyme through loop swapping and hinge variability**. *J Mol Biol* 2011, **411**:143-157.

79. Jones DD: **Recombining low homology, functionally rich regions of bacterial subtilisins by combinatorial fragment exchange**. *PLoS One* 2011, **6**:e24319.

80. Edwards WR, Williams AJ, Morris JL, Baldwin AJ, Allemann RK, Jones DD: **Regulation of beta-lactamase activity by remote binding of heme: functional coupling of unrelated proteins through domain insertion**. *Biochemistry* 2010, **49**:6541-6549.

81. 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.

82. MacDonald JT, Maksimiak K, Sadowski MI, Taylor WR: **De novo backbone scaffolds for protein design**. *Protein Struct Funct Bioinform* 2010, **78**:1311-1325.

83. Correia BE, Ban YEA, Friend DJ, Ellingson K, Xu HY, Boni E, Bradley-Hewitt T, Bruhn-Johannsen JF, Stamatatos L, Strong RK *et al.*: **Computational protein design using flexible backbone remodeling and resurfacing: case studies in structure-based antigen design**. *J Mol Biol* 2011, **405**:284-297.

84. Huang PS, Ban YEA, Richter F, Andre I, Vernon R, Schief WR, Baker D: **RosettaRemodel: a generalized framework for flexible backbone protein design**. *PLoS One* 2011, **6**.

85. Liu W, Hong J, Bevan DR, Zhang YH: **Fast identification of thermostable beta-glucosidase mutants on cellobiose by a novel combinatorial selection/screening approach**. *Biotechnol Bioeng* 2009, **103**:1087-1094.

86. Butz M, Neuenschwander M, Kast P, Hilvert D: **An N-terminal protein degradation tag enables robust selection of highly active enzymes**. *Biochemistry* 2011, **50**:8594-8602.

87. Seelig B: **mRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries**. *Nat Protoc* 2011, **6**:540-552.

88. Kaltenbach M, Hollfelder F: **SNAP display: in vitro protein evolution in microdroplets**. *Methods Mol Biol* 2012, **805**:101-111.

89. Kaltenbach M, Stein V, Hollfelder F: **SNAP dendrimers: multivalent protein display on dendrimer-like DNA for directed evolution**. *Chembiochem* 2011, **12**:2208-2216.

90. Tay Y, Ho C, Droge P, Ghadessy FJ: **Selection of bacteriophage lambda integrases with altered recombination specificity by in vitro compartmentalization**. *Nucleic Acids Res* 2010, **38**:e25.

91. Kornberg A: *For the Love of Enzymes*. Cambridge, MA/London: Harvard University Press; 1989.

92. Yang G, Rich JR, Gilbert M, Wakarchuk WW, Feng Y, Withers SG: **Fluorescence activated cell sorting as a general ultra-high-throughput screening method for directed evolution of glycosyltransferases**. *J Am Chem Soc* 2010, **132**:10570-10577.

93. Liu L, Li Y, Liotta D, Lutz S: **Directed evolution of an orthogonal nucleoside analog kinase via fluorescence-activated cell sorting**. *Nucleic Acids Res* 2009, **37**:4472-4481.

94. Stapleton JA, Swartz JR: **Development of an in vitro compartmentalization screen for high-throughput directed evolution of [FeFe] hydrogenases**. *PLoS One* 2010, **5**:e15275.

95. Hardiman E, Gibbs M, Reeves R, Bergquist P: **Directed evolution of a thermophilic beta-glucosidase for cellulosic bioethanol production**. *Appl Biochem Biotechnol* 2010, **161**:301-312.

96. Varadarajan N, Cantor JR, Georgiou G, Iverson BL: **Construction and flow cytometric screening of targeted enzyme libraries**. *Nat Protoc* 2009, **4**:893-901.

97. Varadarajan N, Pogson M, Georgiou G, Verson BL: **Proteases that can distinguish among different post-translational forms of tyrosine engineered using multicolor flow cytometry**. *J Am Chem Soc* 2009, **131**:18186-18190.

98. Chen I, Dorr BM, Liu DR: **A general strategy for the evolution of bond-forming enzymes using yeast display**. *Proc Natl Acad Sci U S A* 2011, **108**:11399-11404.

99. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret JC, Marquez M, Klibanov AM, Griffiths AD, Weitz DA: **Ultrahigh-throughput screening in drop-based microfluidics for directed evolution**. *Proc Natl Acad Sci U S A* 2010, **107**:4004-4009.

100. Granieri L, Baret JC, Griffiths AD, Merten CA: **High-throughput screening of enzymes by retroviral display using droplet-based microfluidics**. *Chem Biol* 2010, **17**:229-235.

101. Shim J-U, Olguin LF, Whyte G, Scott D, Babtie A, Abell C, Huck WTS, Hollfelder F: **Simultaneous determination of gene expression and enzymatic activity in individual bacterial cells in microdroplet compartments**. *J Am Chem Soc* 2009, **131**:15251-15256.

102. Kintses B, van Vliet LD, Devenish SRA, Hollfelder F: **Microfluidic droplets: new integrated workflows for biological experiments**. *Curr Opin Chem Biol* 2010, **14**:548-555.

Petitioner Merck, Ex. 1141, p. 412