Recent progress in biocatalyst discovery and optimization

Dan E Robertson* and Brian A Steer

The use of enzymes in industrial catalysis continues to grow because of the considerable advantages of natural catalytic systems. The need for enantiomerically pure fine chemicals and the movement away from chemically burdened technologies will drive the acceptance of enzyme-assisted processes. New technologies for enzyme discovery and optimization have enabled the application of enzymes in harsh industrial conditions and in processes demanding stringent selectivity. These discovery and laboratory evolution methods entail genomic approaches that by their nature engender screening of extremely large numbers of gene types and variants. By extension, the fitness of an individual high-throughput screen requires an intelligent, process-targeted assay amenable to a chosen screening platform.

Addresses
Diversa Corporation, 4955 Directors Place, San Diego, CA 92121, USA
*e-mail: drobertson@diversa.com

Current Opinion in Chemical Biology 2004, **8**:141–149

This review comes from a themed issue on
Biocatalysis and biotransformation
Edited by Bernhard Hauer and Stanley M Roberts

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2004.02.010

Abbreviations

DSC	differential scanning calorimetry
E	enantiomeric ratio
eDNA	environmental DNA
ee	enantiomeric excess
EP-PCR	error-prone PCR
FACS	fluorescence-activated cell sorting
GSSM™	gene site saturation mutagenesis™ technology
HT	high throughput
TCP	2,3-dichloro-1-propane
T_m	thermal unfolding transition midpoint
T_{opt}	temperature optimum

Introduction

To access established and emerging markets and provide solutions to industrial issues of specificity, productivity and cost, new as well as established enzyme applications must be enabled or improved with novel, process-optimized biocatalysts [1,2•]. Several enzyme-based processes have been used in industry for years, particularly in textile manufacturing and cellulose degradation [3,4]. The enzymes used in these applications have been derived from a handful of cultured organisms, mainly fungi, and are manufactured at high levels using homologous host expression. Traditional microbial culturing continues to be an important route to new enzyme discovery, although the method is still able to access only a very small percentage of organisms and genomes from narrowly proscribed environmental sources. Many of these enzymes have been optimized using directed mutagenesis technologies. These efforts have recently been reviewed [5–8]. This review concentrates on developments in the years 2002–2003.

Enzyme discovery using novel genomic approaches

Virtually all terrestrial, aquatic and marine, as well as human-impacted environments, support microbial populations participating in complex and often industrially interesting chemical cycles under extant conditions of temperature, pH, salinity, pressure and environmental chemical concentrations. Capturing the genes of organisms that have evolved as participants in these biotopes promises to revolutionize and broaden enzyme applications in the chemical industry. Key to this gene capture are new culturing and genomic technologies coupled with chemistry-specific high-throughput (HT) expression assays or gene sequence hybridization methods [9–11].

HT culturing is a novel technology designed to expand the number of cultured genomes and the spectrum of industrially useful enzymes [12]. Beginning with a complex consortium of microorganisms collected from a targeted biotope, the method compartmentalizes the individual microorganisms in gel microdroplets, places the microdroplets into a glass column and passes media, specific to the biotope of origin, through the array. The technique thereby mimics the original environment by allowing interspecies interactions while growing out colonies of individual organisms. The individual colonies are isolated by fluorescence-activated cell sorting (FACS) of the gel microdroplets. This method promises to reveal new species from interesting biotopes and to provide DNA for capture of new biocatalysts.

Genomic methods for broadening the spectrum of available enzymes have also been developed that enable gene discovery directly from any environment [9]. The approach directly extracts DNA, fractionating it by size and clones the fragments into a vector designed for host transformation [13]. The environmental DNA (eDNA) clones generated can be arrayed into a chosen screening format and subjected to assays targeted to discovery of particular gene sequences or gene product activities. Once an activity or gene sequence is detected, the DNA fragment of the isolated clone is sequenced. This gene can be subcloned and expressed in a chosen host system. In addition to the comprehensive scope of the

eDNA libraries, they have broad utility in that they can be used repeatedly for expression or sequence-based screening using any assay format. They may be stored indefinitely, normalized to gain uniform representation of all genomes or enriched for particular sequence homologues.

Genomic methods are useful for comprehensive genome collection, but the techniques generate 10^5 to 10^9 clones requiring expression or sequence screening. This number is often too large for conventional analytical assays (e.g. HPLC and GC), although we will note creative exceptions. Critical to the discovery of desirable enzymes, then, is a rapid discovery assay that correctly and directly targets the genes of interest. There are several assay paradigms that may be used [14].

One overall approach is a tiered strategy that begins with a general screen, often with a surrogate substrate, to detect a general chemistry carried out by the enzyme class of interest, or with a sequence probe designed to be broadly homologous to the sequence of the class. This step, which can narrow the scope of clones, can be followed by more specific assays for aspects of enzyme phenotype applicable to the targeted process.

The simplest and most comprehensive approach to library screening is direct selection. Selections can access the largest numbers of expressing entities, essentially an infinite range due to the lack of necessity for arraying and sequentially detecting the results of an enzymatic reaction. Development of a selection requires recruitment of a necessary or a detrimental activity for clone growth.

One selection, relying upon the necessity for medium nitrogen has been used for isolation of nitrilase enzymes from eDNA libraries [11,15]. Over 650 eDNA libraries were screened and nitrilase clones were identified by growth, due to a positive clone's ability to hydrolyze a nitrile substrate into carboxylic acid and ammonia products for nitrogen generation. More than 200 unique nitrilase sequences were discovered using this assay and 137 of these were further cloned, subjected to phylogenetic analysis and characterized in subsequent assays for enantioselectivity towards three industrially interesting substrates. The results of this study were interesting in that the nitrilases discovered by selection on surrogate substrates sorted into five distinct sequence clades and the chemospecificity and enantiospecificity sorted, in large part, into unique clade groupings.

Various formats may be used for high-density screening. Microtiter formats range from 96 wells to 1536 wells, although even at the highest commercially available densities only around 10^5 to 10^6 assay events may be completed per day. FACS has been adapted for screening eDNA libraries achieving throughput levels of 10^9 events per day [16].

A new ultra high density screening format, called GigaMatrix^{TM} technology has been developed that incorporates over 100 000 wells in a microtiter-sized footprint [17••,18,19]. This system sequesters individual particles (i.e. cells, clones, proteins, etc.) in through-hole wells that can be incubated and subsequently screened using fluorescence detection via imaging by a cooled CCD camera. A retrieval device automatically removes positives and deposits them into addressed microtiter wells. The system has been used successfully for discovery of enzymes from eDNA libraries.

Probably the most universal assay type used in enzyme discovery is direct expression screening using substrates with bonds attacked by the enzyme class of interest. Detection can be by color change, fluorescence intensity or wavelength change, polarization, resonance energy transfer or any other phenomenon easily detectable in a screening format. A recent review has surveyed the journal literature on this subject up through 2001 [14]. Because bond attack in substituted surrogates may not define the ultimate enzyme of interest, this method of screening usually requires tiered screening to first identify a group of enzyme candidates followed by more specific screening using low-throughput methods.

One consideration when screening clones in microtiter format or other modes is situations where substrate solubility, stability or optical response are critical or where cell materials or media may add significant background. Stability of the substrate is an issue, particularly for nitrophenyl and umbelliferyl esters. Leroy *et al.* [20] have surveyed the use of acyloxymethyl ethers of umbelliferone for hydrolysis of esters by lipases and esterases. Testing solubility and stability of a range of substituted variants, the authors conclude that isobutyrate- and pivaloxymethyl-ether substituents are optimal for general esterase activity screening.

Hydrolytic enzymes have often been assayed by observing the pH change accompanying hydronium ion release during the reaction. Many pH-sensitive chromogenic and fluorogenic substrates are available and their application to esterase and dehalogenase discovery and optimization have been reviewed [21]. Using pH detection by bromophenol blue in a microtiter-based screen, Gray *et al.* [22] screened a library of dehalogenase clones for thermostability and enhanced turnover and constructed a gene producing an enzyme with significantly enhanced refolding and stability after temperature challenge.

Coupled assays, wherein an indirect reaction with product generates an optical response, have also been used to detect enzymatic reactions. A chemosensor system, wherein calcein is used as a sensor for amino acids, has been used as a sensitive assay for protease activity [23]. Mateo *et al.* [24] have reported the use of periodate

oxidation of diols to quantitatively monitor the kinetics of epoxide hydrolase activity.

Finally, sequence-based screening methods have been used for capture of gene homologues from DNA libraries. One key method has been PCR-based hybridization. In a technique called biopanning, which is designed to specifically enrich for gene sequence classes, haloalkane dehalogenase homologues were targeted using a set of degenerate primers that were used to prescreen eDNA libraries derived from various soil biotopes [10]. The primers incorporating gene fragments with two internally conserved sequence motifs from three known dehalogenases were used as probes to pull out full-length sequences with significant identity and with dehalogenase activity. Screening of 461 eDNA libraries resulted in capture of >70 unique clones greater than 400 bp in length. Ultimately, 19 novel full-length dehalogenase genes were cloned and sorted into a phylogeny and characterized with respect to their specificity for industrially important substrates, trichloropropane and dibromoethane.

Enzyme optimization by directed mutagenesis

Directed mutagenesis techniques have been used with great success for the diversification of gene sequences and optimization of enzyme phenotypes. Since the beginning of 2002, several reviews have appeared describing random and rational approaches to *in vitro*, *in vivo* and *in silico* codon variation and gene recombination for generation of gene variant libraries [5–8,25,26]. These reviews focus on the methods for mutant library generation and on the results achieved in enzyme phenotype improvement. Other patented technologies have also appeared that create comprehensive directed mutagenesis libraries. Gene site saturation mutagenesis™ technology (GSSM) creates a library with representatives of all possible codon variants in a gene [27•,28,30,31•,51•]. GeneReassembly™ technology recombines several parental gene sequences, with control over demarcation points and without regard for homology, to create a purely stochastic progeny library of gene variants [32•,33,34]. By virtue of the comprehensive nature of these various techniques used for searching gene sequence space, the size of libraries generated for screening is large, often requiring 10^5–10^{10} screening events for full library coverage. Because of library size and because a directed mutagenesis effort is undertaken to optimize a very specific aspect of enzyme phenotype, optimal assay development is critical to a successful outcome, that is, a gene product with optimal fitness.

Enzyme specificity

Current industrial chemistry struggles to provide efficient processes for oxidations. Specificity and economics of current processes plus the demand for new chemistries can drive the acceptance of biocatalysts in this key area. Directed mutagenesis of monooxygenases and dioxygenases to enhance activity, alter specificity and redirect reducing equivalents has been a focus of several laboratories whose efforts have been reviewed by Cirino and Arnold [35]. More recently, the P450 BM-3 monooxygenase from *Bacillus megaterium*, normally specific for medium-chain fatty acids, has been evolved to accept small hydrocarbon substrates such as propane and propylene and to function at high rates [36]. The regioselectivity and enantioselectivity of P450-mediated oxidation of long chain alkanes has also been improved by directed mutagenesis of this same enzyme and the oxidation rate for propane has been enhanced beyond even the best variant found earlier [37]. These optimizations were achieved using error-prone PCR (EP-PCR) and gene shuffling and colorimetric assays for octane hydroxylation involving *p*-nitrophenoxy octane or its octyl ether.

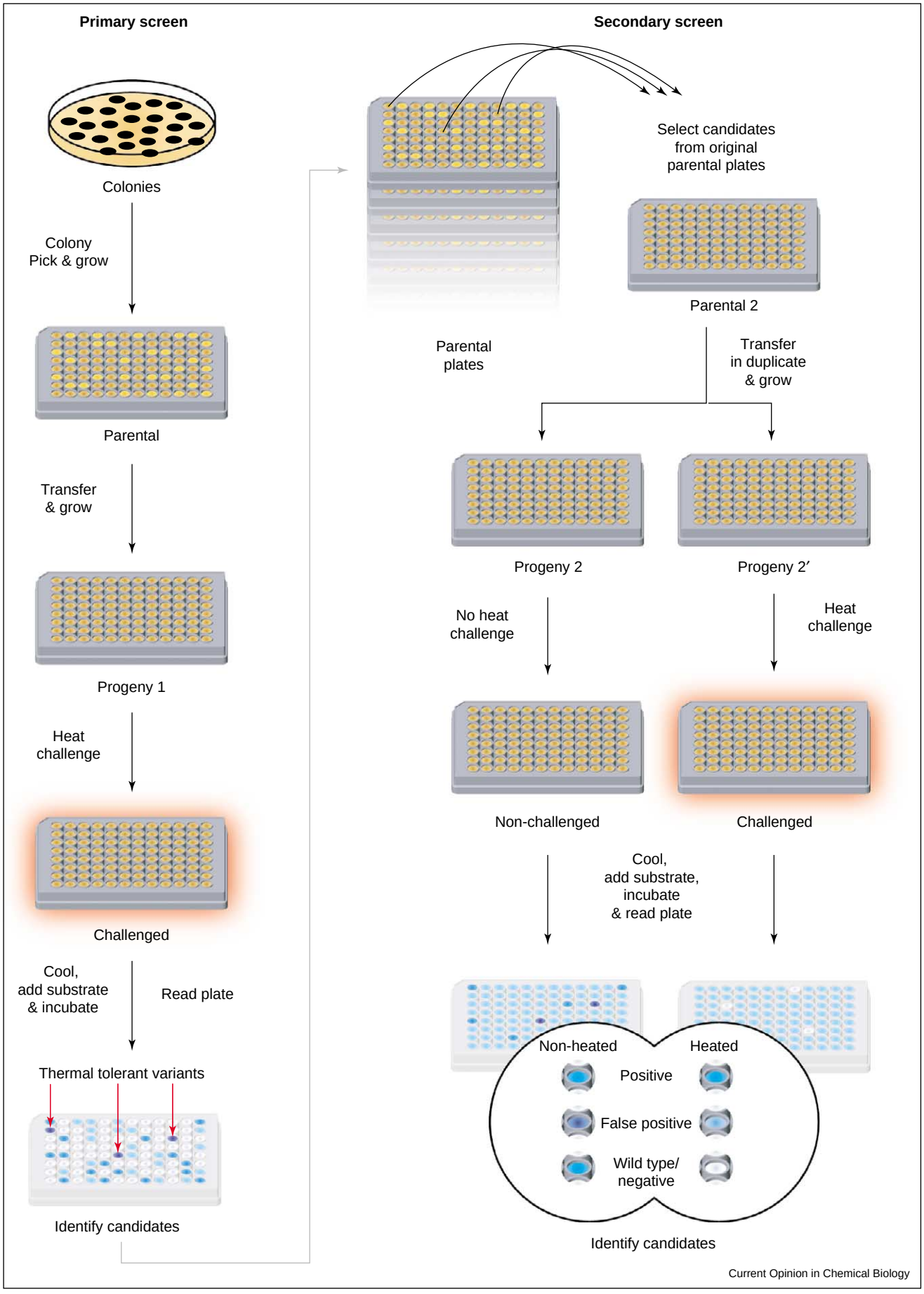
Rate increases in nerve agent and pesticide degrading enzymes have been targeted recently using display technologies. Cho *et al.* [38] performed EP-PCR and gene shuffling on the *Pseudomonas diminuta* MG organophosphohydrolase. Using ice nucleation protein-linked surface display on *Escherichia coli* followed by a solid-phase colorimetric assay on methyl parathion, mutations were detected giving a 25-fold increase in turnover. Griffiths and Tawfik [39] developed a novel display and compartmentation paradigm using microbeads and an assay scheme involving hydrolysis of a biotinylated paraoxon derivative coupled to flow cytometric fluorescence detection of an antibody–product interaction. This assay format yielded a mutant enzyme with a k_{cat} 63-fold higher than its wild-type parent.

Using EP-PCR and DNA shuffling, a haloalkane dehalogenase from *Rhodococcus* was optimized for hydrolysis of 2,3-dichloro-1-propane (TCP) and subsequently cloned into a TCP-utilizing strain of *Agrobacterium radiobacter* to create a strain capable of using TCP as sole carbon source [40•]. The assay used to isolate up-mutants involved colorimetric detection of hydrolysis-associated hydronium ion production on solid agar medium. This strategy of evolution followed by incorporation into a metabolically compatible host can enable remediation schemes for toxic chemicals.

Synthesis of cephalosporins can be enhanced enzymatically by reducing the number of steps in the chemical synthesis. Using EP-PCR and shuffling of a glutaryl acylase gene from *Pseudomonas* coupled with a fluorescamine-based assay for determination of the ratio of k_{cat}/K_M for adipyl-7-aminocephalosporanic acid to that for glutaryl-7-aminocephalosporanic acid, mutants were isolated that had up to 10-fold increases in catalytic efficiency for the desired adipyl substrate [41].

Enantioselectivity

One aspect of enzyme phenotype of particular significance to the specialty chemical and pharmaceuticals


Primary screen
Secondary screen
Colonies
Colony
Pick & grow
Parental
Transfer
& grow
Progeny 1
Heat
challenge
Challenged
Cool,
add substrate
& incubate
Read plate
Thermal tolerant variants
Identify candidates
Select candidates
from original
parental plates
Parental 2
Parental
plates
Transfer
in duplicate
& grow
Progeny 2
Progeny 2′
No heat
challenge
Heat
challenge
Non-challenged
Challenged
Cool,
add substrate,
incubate
& read plate
Non-heated
Heated
Positive
False positive
Wild type/
negative
Identify candidates
Current Opinion in Chemical Biology

sectors is enantioselectivity. Various hydrolases, oxidoreductases, transferases and isomerases have been viewed as potential candidates to perform chiral resolutions and syntheses. Recent, comprehensive reviews have focused on directed mutagenesis methods applied to the problem, cited examples and highlighted the basic challenge for enantioselective screens; the discrimination of enantiomers in an HT detection modality [42–45].

Quick E is a colorimetric screening method developed by Kazlauskas that has been used successfully to identify increases in hydrolase enantioselectivity. Quick E individually monitors the rate of hydrolysis of each enantiomer, using a pH indicator, in the presence of a substrate reference compound. From the two hydrolysis rates, a value for enantioselectivity can be calculated. Recently, the technique was used to screen directed mutagenesis mutants of an esterase from *Pseudomonas fluorescens* and to identify mutations in specific residues that lead to significant increases in enantioselectivity for two ester substrates [46]. Saturation mutagenesis at these residues led to selectivity increases for methyl bromo-2-methylpropanoate (E increased from 12 to 21) and ethyl-3-phenylbutanoate (E increased from 3.7 to 12).

In a directed mutagenesis effort generating simultaneous variations in four binding site residues, Koga *et al.* [29•] inverted the enantioselectivity of a lipase from *Burkholderia cepacia* using a technique called SIMPLEX (single molecule-PCR-linked *in vitro* expression). The residues were chosen by viewing the enzyme crystal structure and choosing residues near the acyl-binding region. The residues were combinatorially substituted with seven hydrophobic counterparts, creating a library of 2400 variants. HPLC screening for hydrolysis of a racemic mixture of *p*-nitrophenyl 3-phenylbutyrate detected several combinations that reversed enantioselectivity.

Rather than assay the outcome of evolutionary efforts aimed at optimizing the selectivity of an enzyme for an unnatural isomer, directed mutagenesis of aldolases aims to alter the stereochemistry of C–C bond formation. DNA shuffling following EP-PCR was applied to the Class II FBP aldolase of *Escherichia coli* to shift its specificity towards fructose-1,6-bisphosphate, resulting in a 100-fold increase in stereospecificity [47]. Wada *et al.* [48], using similar directed mutagenesis methodology, effected a reversal of enantioselectivity in the Neu5Ac aldolase of *E. coli*. Both studies used a coupled assay in microtiter format, which measured NADH levels.

Many investigators, faced with a large library to screen, have developed HT variations of conventional separation and detection chemistries such as IR-thermography, capillary electrophoresis and mass spectrometry. References to these applications have been noted recently [49]. Mass spectrometry is particularly useful for discrimination of enantiomers when one enantiomer can be isotopically labeled. This substitution methodology is benign with regard to its effect on the enzyme specificity and the technique can be designed to sample microtiter formats at reasonably high throughput. Electrospray ionization MS was used to iteratively screen *Bacillus subtilis* lipase directed mutagenesis libraries generated by EP-PCR [50]. In total, 27 000 mutants were screened using deuterium-labeled pseudo-meso-1,4-diacetoxy-2-cyclopentene, resulting in the identification of several mutants with significant enantiomeric excess (*ee*) increases. In an optimization effort using a nitrilase discovered from eDNA, DeSantis *et al.* [51•] screened >31 000 clones from a GSSM^{TM} library [26] for increased *ee* for (*R*)-4-cyano-3-hydroxybutyric acid. In this case, the (*R*)-enantiomer of the 3-hydroxyglutarylnitrile substrate was N^{15}-enriched. GSSM^{TM} technology, which saturated each position in the protein with the 19 side chain variants, coupled with the electrospray-MS screen yielded a two-base change in a single codon, which increased the *ee* of the nitrilase to from 88% to 99% *ee* at 3M substrate concentration.

Other novel approaches to enantioselective assay have been reported recently. Tielmann *et al.* [49] report the use of Fourier transform infrared spectroscopy (FT-IR) for HT screening rates approaching 10 000 events per day. As in MS techniques, FT-IR requires isotopic labeling of one enantiomer to give a pseudoenantiomeric substrate mixture. Another interesting and less capital-intensive assay method uses monoclonal antibodies raised to one enantiomer of a racemic substrate or product mixture. The specific antibody can then be used for ELISA assay in a microtiter format. This concept was demonstrated by Taran *et al.* [52•] using an antibody raised to (*S*)-mandelic acid. The antibody had cross-reactivity to the (*R*)-enantiomer of less than 1%.

Thermostability

Many industrial applications of enzymes require that the enzyme function at, or survive exposure to, elevated

(Figure 1 Legend) Microtiter plate screening for thermal tolerance improvements. In the primary screen, the library clones are arrayed in microtiter plates by colony picking into liquid medium. Differences in colony size or the amount of bacterial cells transferred during picking lead to uneven growth in the parental plate. A liquid–liquid inoculation into a second plate results in much more uniform growth (progeny plates) and reserves the parental plate for storage. The progeny plates are then subject to a heat challenge followed by the measurement of residual activity at mesophilic temperature. In the secondary screen, the primary candidates are selected from the reserve of parental plates and inoculated into a new plate to generate a second parental plate for the secondary screen. Two identical progeny plates are made to allow the comparison of enzyme activity with and without a heat challenge for the identification of false positives. False positives will generally show greater activity in the non-heat-challenged plate, and residual activity in the heated plate due to elevated levels of enzyme.

temperatures [3,4,6,53]. Although many of the types of interactions that play a role in protein thermal stability are known [54] the relative contributions of these interactions in conferring thermal stability for any given protein are poorly understood and are not universal. Thus, directed mutagenesis is a key method for the improvement of enzyme thermal tolerance as it allows the random selection of mutations that improve thermal tolerance. Its fundamental advantage over rational approaches is that it forces the enzyme to find the best solution for surviving exposure to heat. The solution might be increased thermostability, improved folding reversibility or a combination of the two and it is characteristic to a given enzyme and rarely predictable.

The key factor in the success of a directed mutagenesis endeavor is the screening assay used to select improved variants. In most cases, assays are designed to select for variants with better tolerance to heat rather than increased activity at elevated temperature.

Challenging screens for thermal tolerance have resulted in the identification of enzyme variants with substantially improved thermostability [30,31•,55•]. In these studies, the improvement in thermostability was measured using differential scanning calorimetry (DSC) rather than by the measurement of relative activity with increasing temperature. Increases in the transition midpoint for thermal unfolding (T_m) from 10°C [55•] to 12°C [30] to as high as 35°C [31•] were achieved. Increases in reaction temperature optima (T_{opt}) are often higher (most likely due to substrate stabilization effects) than shifts in T_m as measured by techniques such as DSC that directly measure thermostability. For example, directed mutagenesis to enhance the thermal resistance of an amylase resulted in an increase in T_m of 10°C but a 15°C shift in T_{opt} [55•].

Although the terms 'thermal tolerance' and 'thermostability' are often used interchangeably, successful screening assays for thermal tolerance can identify mutations that do not enhance thermostability but, rather, effect changes in protein folding dynamics. A thermal tolerant variant is not necessarily thermostable because improving the reversibility of thermal denaturation can give rise to thermal tolerance rather than increased thermostability [56••]. In some cases, combinations of mutations that augment both phenotypes are observed [22].

In general, there are two main approaches that have been used for high-throughput screening for variants with increased residual activity. The first approach makes use of agar plates or membranes and in the second approach the assays are performed in microtiter plates. The main benefit of the agar/membrane-based approach is that several thousand variants from the mutagenic library can be assayed per plate. (Most HT screens in microtiter plates utilize 96- [31•,56••,57,58] or 384-well [22,30] plates with one variant per well.) It is also a time-saving feature that clones can be transformed, plated and the colonies directly replicated to an assay plate or membrane. Screening with membranes has been used recently to improve the thermal tolerance of an amino acid oxidase [59] and an amylase [55•]. Filter lifts from agar plates were used effectively to enhance the stability of glucose dehydrogenase [60] and *N*-carbanyl-D-amino acid amidohydrolase [61] and an agar plate screen employing clearing zone detection was applied to improve glucanase thermal tolerance [62].

Microtiter plate assays can be highly quantitative and they allow kinetic measurement of enzyme activities. In most cases, turnover is quantified by absorbance or fluorescence using model substrates such as dye-linked polysaccharides, *p*-nitrophenyl or methylumbelliferyl substrate derivatives [30,56••,58] or colorimetric indicators [22,63].

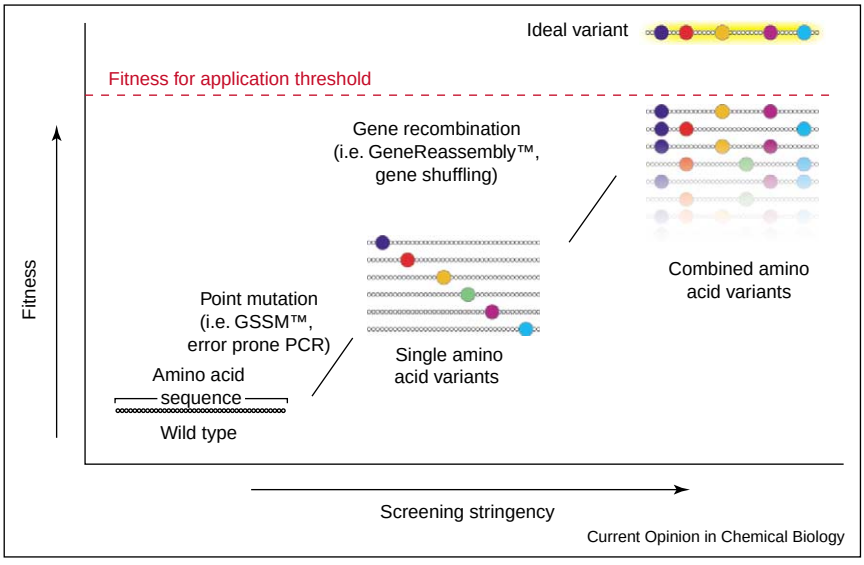
One aspect of screening for thermal tolerance that adds complication is the variability of enzyme concentration. Variants within a library may have variable expression levels and affect cell density and colony size (in the case of agar plate/membrane screening), which can all influence the amount of each individual enzyme variant tested in the screen. As such, to eliminate false positives, the comparison of heated and non-heated samples is essential in the identification of truly improved variants [30,31•,56••,63]. This comparison can be made in the primary screen or in a secondary screen where positive mutants identified in the primary screen are re-tested and compared under heated and non-heated conditions (Figure 1).

The final mutations identified in the secondary screen that impart thermal tolerance can then be recombined additively, using mutagenesis or combinatorially using shuffling or Gene Reassembly [31•,32•,58,61]. The combinatorial library is screened using the same screen (Figure 1) but with much greater stringency. In a screen for thermal tolerance, the increased stringency would be a heat challenge at a higher temperature or for a longer duration. Increasing the screening stringency allows the identification of the fittest variant from the combinatorial library of individual mutants (Figure 2).

Expression

Homologous and, more importantly for genes derived from cDNA or directed mutagenesis libraries, heterologous expression of discovered and optimized enzymes is often an issue for the economic implementation of an enzyme. Transcription and translation rates and levels and protein solubility are factors that can also be optimized using directed mutagenesis and associated assays. Waldo [64] has reviewed the literature and emphasized

Figure 2



Improving enzyme fitness with increasing screening stringency. Fitness may be defined in any screen as the best combination of phenotypic variables, for example, thermostability and rate of enzyme turnover. The primary screen identifies single amino acid mutations that improve fitness. Unlike EP-PCR, GSSM has an advantage in that all possible amino acid side chain variations are parsed with equal probability. The single mutations may be built into the gene by site-directed mutagenesis or methods such as shuffling and gene reassembly may be used for combinatorial reassembly. After the primary single amino acid mutations are recombined, the library is screened at higher stringency, selecting those variants with even greater fitness. Ultimately the variant with the highest fitness is obtained.

the use of fusion reporter tags and immunological detection as function independent means to report expression and solubility of proteins. Heterologous expression of a fungal laccase in *Saccharomyces* host was improved using a combination of directed mutagenesis methods including EP-PCR, *in vivo* recombination and staggered extension in conjunction with direct activity screens to increase expression, k_{cat} and stability of the enzyme [65]. Gene reassembly of three parental genes discovered from eDNA libraries was used to engineer an enhanced α-amylase with optimal activity at pH 4.5 and high levels of heterologous expression in a *Pseudomonas* host system [34].

Conclusions

Biocatalytic technologies will ultimately gain universal acceptance when enzymes are perceived to be robust, specific and inexpensive (i.e. process compatible). Genomics-based gene discovery from novel biotopes and the broad use of technologies for accelerated laboratory evolution promise to revolutionize industrial catalysis by providing highly selective, robust enzymes. As HT screening technologies, ever more intelligent assay design and process conscious research are brought to bear, the utility and economics of enzymes should mandate the use of these mild, selective catalysts.

Acknowledgements

We thank Michael Kulwiec for figure design and illustration.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. DeSantis G: **Enzymes in organic synthesis**. *Mod Drug Discovery* 2003, **5**:43-47.

2. • Schoemaker HE, Mink D, Wubbolts MG: **Dispelling the myths–biocatalysis in industrial synthesis**. *Science* 2003, **299**:1694-1697.

A recent, comprehensive overview of the field of biocatalysis including the successes and, most importantly, the specific chemical challenges which may be addressed by enzyme catalysis.

3. Kirk O, Borchert TV, Fuglsang CC: **Industrial enzyme applications**. *Curr Opin Biotechnol* 2002, **13**:345-351.

4. van Beilen JB, Li Z: **Enzyme technology: an overview**. *Curr Opin Biotechnol* 2002, **13**:338-344.

5. Turner NJ: **Directed evolution of enzymes for applied biocatalysis**. *Trends Biotechnol* 2003, **21**:474-478.

6. Cherry JR, Fidantsef AL: **Directed evolution of industrial enzymes: an update**. *Curr Opin Biotechnol* 2003, **14**:438-443.

7. Tao H, Cornish VW: **Milestones in directed enzyme evolution**. *Curr Opin Chem Biol* 2002, **6**:858-864.

8. Dalby PA: **Optimising enzyme function by directed evolution**. *Curr Opin Struct Biol* 2003, **13**:500-505.

9. Short JM: **Recombinant approaches for accessing biodiversity**. *Nat Biotechnol* 1997, **15**:1322-1323.

10. Gray KA, Richardson TH, Robertson DE, Swanson PE, Subramanian MV: **Soil-based gene discovery: a new technology to accelerate and broaden biocatalytic applications**. *Adv Appl Microbiol* 2003, **52**:1-27.

11. Robertson DE, Chaplin JA, DeSantis G, Podar M, Madden M, Chi E, Richardson T, Milan A, Miller M, Weiner DP *et al.*: **Exploring**

nitrilase sequence space for enantioselective catalysts. *Appl Environ Microbiol* 2004, **70**:in press.

12. Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M: **Cultivating the uncultured**. *Proc Natl Acad Sci USA* 2002, **99**:15681-15686.

13. Short JM: **Gene expression library produced from DNA from uncultivated microorganisms and methods for making the same**. US patent 2001, US6280926.

14. Wahler D, Reymond JL: **High-throughput screening for biocatalysts**. *Curr Opin Biotechnol* 2001, **12**:535-544.

15. DeSantis G, Zhu Z, Greenberg WA, Wong K, Chaplin J, Hanson SR, Farwell B, Nicholson LW, Rand CL, Weiner DP *et al.*: **An enzyme library approach to biocatalysis: development of nitrilases for enantioselective production of carboxylic acid derivatives**. *J Am Chem Soc* 2002, **124**:9024-9025.

16. Short JM, Keller M: **High throughput screening for novel enzymes**. US patent 2001, US6174673.

17. •• Lafferty M, Dycaico M: **GigaMatrix™-a novel ultra high throughput protein optimization and discovery platform**. *Methods Enzymol* 2004, **388**:in press.
This reference details the development of GigaMatrix™ technology, an ultra-high density screening format, amenable to fluorescence-based expression screening. The technology enables microtiter-style screening at a level of 10^9 events/day.

18. Short JM, Keller M, Lafferty WM: **High throughput or capillary-based screening for a bioactivity or biomolecule**. US patent application 2003, S20030049841A1.

19. Lafferty M: **Capillary array-based sample screening**. US patent application 2002, 20020015997A1.

20. LeRoy E, Bensel N, Reymond J-L: **A low background high throughput screening (HTS) fluorescence assay for lipases and esterases using acyloxymethylethers of umbelliferone**. *Bioorg Med Chem Lett* 2003, **13**:2105-2108.

21. Griswold KE: **pH sensing agar plate assays for esterolytic enzyme activity**. In *Methods in Molecular Biology*, vol 230. Edited by Arnold FH, Georgiou G. Totowa: Humana Press; 2003:203-211.

22. Gray KA, Richardson TH, Kretz K, Short JM, Bartnek F, Knowles R, Kan L, Swanson PE, Robertson DE: **Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase**. *Adv Synth Catal* 2001, **343**:607-617.

23. Dean KES, Klein G, Renaudet O, Reymond J-L: **A green fluorescent chemosensor for amino acids provides a versatile high-throughput screening (HTS) assay for proteases**. *Bioorg Med Chem Lett* 2003, **13**:1653-1656.

24. Mateo C, Archelas A, Furstoss R: **A spectrophotometric assay for measuring and detecting an epoxide hydrolase activity**. *Anal Biochem* 2003, **314**:135-141.

25. Zhao H, Chockalingam K, Chen Z: **Directed evolution of enzymes and pathways for industrial biocatalysis**. *Curr Opin Biotechnol* 2002, **13**:104-110.

26. Kretz K, Richardson T, Gray KA, Robertson DE, Tan X, Short JM: **Gene site saturation mutagenesis (GSSM): a comprehensive mutagenesis technique**. *Methods Enzymol* 2004, **388**:in press.

27. • Short JM: **Saturation mutagenesis in directed evolution**. US patent 2001, US6171820.
This patent describes a comprehensive codon mutagenesis technology that enables the alteration of every codon in a gene to each of its 63 alternatives, hence allowing access to all 19 amino acid side chain variants at each residue in a protein sequence. The method surpasses other error-based techniques such as EP-PCR that can only achieve between five and seven side-chain variations and only one base alteration per codon.

28. Short JM: **Saturation mutagenesis in directed evolution**. US patent 2001, 6,562,594.

29. • Koga Y, Kato K, Nakano H, Yamane T: **Inverting enantioselectivity of Burkholderia cepacia KWI-56 lipase by combinatorial mutation and high-throughput screening using single-molecule PCR and *in vitro* expression**. *J Mol Biol* 2003, **331**:585-592.
The authors describe a mutagenesis technique, employing single-molecule DNA amplification followed by *in vitro* expression, designed to create combinatorial mutagenesis libraries. The authors demonstrate an application targeting the switch in enantioselectivity for a lipase.

30. Garrett J, Kretz K, O'Donoghue E, Kerovuo J, Kim W, Barton N, Hazlewood GP, Short JM, Robertson DE, Gray K: **Enhancing the thermal tolerance and gastric performance of a microbial phytase for use as a phosphate-mobilizing monogastric feed supplement**. *Appl Environ Microbiol* 2004, in press.

31. • Palackal N, Brennan Y, Callen WN, Dupree P, Frey G, Goubet F, Hazlewood GP, Healey S, Kang YE, Kretz KA *et al.*: **An evolutionary route to xylanase process fitness**. *Protein Sci* 2004, **13**:494-503.
A microtiter plate screen was used to identify mutations in a xylanase that increased thermostability. The primary mutants were reassembled combinatorially and with a second screen at higher stringency, variants were identified with remarkable increases in T_m relative to wild-type ranging from 28–35°C.

32. • Short JM: **Synthetic ligation reassembly in directed evolution**. US patent 2003, US6537776.
A patent describing a non-PCR, non-amplification-based technology for the reassembly of gene fragments. The virtue of the approach lies in its ability to use parental genes without regard to their relative sequence homology and to choose demarcation points in the parental sequences at will. The resulting library has a uniform statistical distribution of parental fragments, with the complete parental sequences recurring at equal frequency to any progeny.

33. Short JM: **Synthetic ligation reassembly in directed evolution**. US patent 2003, US6605449.

34. Richardson TH, Tan X, Frey G, Callen W, Cabell M, Lam D, Macomber J, Short JM, Robertson DE, Miller C: **A novel, high performance enzyme for starch liquefaction. Discovery and optimization of a low pH, thermostable alpha-amylase**. *J Biol Chem* 2002, **277**:26501-26507.

35. Cirino PC, Arnold FH: **Protein engineering of oxygenases for biocatalysis**. *Curr Opin Chem Biol* 2002, **6**:130-135.

36. Glieder A, Farinas ET, Arnold FH: **Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase**. *Nat Biotechnol* 2002, **20**:1135-1139.

37. Peters MW, Meinhold P, Glieder A, Arnold FH: **Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3**. *J Am Chem Soc* 2003, **125**:13442-13450.

38. Cho CM, Mulchandani A, Chen W: **Bacterial cell surface display of organophosphorus hydrolase for selective screening of improved hydrolysis of organophosphate nerve agents**. *Appl Environ Microbiol* 2002, **68**:2026-2030.

39. Griffiths AD, Tawfik DS: **Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization**. *EMBO J* 2003, **22**:24-35.

40. • Bosma T, Damborsky J, Stucki G, Janssen DB: **Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene**. *Appl Environ Microbiol* 2002, **68**:3582-3587.
An integrated approach to bioremediation wherein an enzyme was optimized for activity on the first substrate in a metabolic pathway, TCP, and then cloned into a host that utilized the enzyme's product, rendering the resultant host capable of using TCP as sole carbon source.

41. Otten LG, Sio CF, Vrielink J, Cool RH, Quax WJ: **Altering the substrate specificity of cephalosporin acylase by directed evolution of the beta subunit**. *J Biol Chem* 2002, **277**:42121-42127.

42. Reetz MT: **New methods for the high-throughput screening of enantioselective catalysts and biocatalysts**. *Angew Chem Int Ed Engl* 2002, **41**:1335-1338.

43. Reetz MT: **An Overview of high-throughput screening systems for enantioselective enzymatic transformations**. In *Methods in Molecular Biology*, vol 230. Edited by Arnold FH, Giorgiou G. Totowa: Humana Press; 2003:259-282.

44. Bornscheuer UT, Bessler C, Srinivas R, Krishna SH: **Optimizing lipases and related enzymes for efficient application**. *Trends Biotechnol* 2002, **20**:433-437.

45. Bornscheuer UT: **Methods to increase enantioselectivity of lipases and esterases**. *Curr Opin Biotechnol* 2002, **13**:543-547.

46. Horsman GP, Liu AMF, Henke E, Bornscheuer UT, Kazlauskas RJ: **Mutations in distant residues moderately increase the enantioselectivity of Pseudomonas fluorescens esterase towards methyl 3-bromo-2-methylpropionate and ethyl 3-phenylbutyrate**. *Chem Eur J* 2003, **9**:1933-1939.

47. Williams GJ, Domann S, Nelson A, Berry A: **Modifying the stereochemistry of an enzyme-catalyzed reaction by directed evolution**. *Proc Natl Acad Sci USA* 2003, **100**:3143-3148.

48. Wada M, Hsu C-C, Franke D, Mitchell M, Heine A, Wilson I, Wong C-H: **Directed evolution of N-acetylneuraminic acid aldolase to catalyze enantiomeric aldol reactions**. *Bioorg Med Chem Lett* 2003, **11**:2091-2098.

49. Tielmann P, Boese M, Luft M, Reetz MT: **A practical high-throughput screening system for enantioselectivity by using FTIR spectroscopy**. *Chemistry* 2003, **9**:3882-3887.

50. Funke SA, Eipper A, Reetz MT, Otte N, Thiel W, Van Pouderoyen G, Dijkstra BW, Jaeger K-E, Eggert T: **Directed evolution of an enantioselective *Bacillus subtilis* lipase**. *Biocatal Biotransform* 2003, **21**:67-73.

51. • DeSantis G, Wong K, Farwell B, Chatman K, Zhu Z, Tomlinson G, Huang H, Tan X, Bibbs L, Chen P *et al.*: **Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM)**. *J Am Chem Soc* 2003, **125**:11476-11477.
A unified application of discovery and evolution targeted to the solution of a specific problem of process-scale enantioselective catalysis; a newly discovered nitrilase was comprehensively mutated and screened using LC-MS to achieve high *ee* at high substrate concentrations. The essential mutation required two base changes in a single codon, achievable only by GSSM technology.

52. • Taran F, Gauchet C, Mohar B, Meunier S, Valleix A, Renard PY, Cremonon C, Grassi J, Wagner A, Mioskowski C: **High-throughput screening of enantioselective catalysts by immunoassay**. *Angew Chem Int Ed Engl* 2002, **41**:124-127.
A novel use of monoclonal antibody-ELISA techniques designed for the HT enantioselective assay of an enzyme library. The authors make the point that analytical assays for enantioselectivity are capital intensive and that this assay paradigm is affordable and specific.

53. van Den BB: **Extremophiles as a source for novel enzymes**. *Curr Opin Microbiol* 2003, **6**:213-218.

54. Yano JK, Poulos TL: **New understandings of thermostable and peizostable enzymes**. *Curr Opin Biotechnol* 2003, **14**:360-365.

55. • Kim YW, Choi JH, Kim JW, Park C, Kim JW, Cha H, Lee SB, Oh BH, Moon TW, Park KH: **Directed evolution of Thermus maltogenic amylase toward enhanced thermal resistance**. *Appl Environ Microbiol* 2003, **69**:4866-4874.
This work is an example of a solid improvement in thermostability where an amylase variant with a T_m of 87°C was identified using a filter lift screening assay. The T_m was increased by 11°C even though the wild-type parent already possessed considerable thermostability ($T_m = 76$°C).

56. •• Zhang N, Suen WC, Windsor W, Xiao L, Madison V, Zaks A: **Improving tolerance of *Candida antarctica* lipase B towards irreversible thermal inactivation through directed evolution**. *Protein Eng* 2003, **16**:599-605.
In this study, the authors used a microtiter plate screen for thermal tolerance and identified two lipase mutants that enhanced thermal tolerance by improving the reversibility of folding.

57. Cirino PC, Georgescu R: **Screening for thermostability**. *Methods Mol Biol* 2003, **230**:117-125.

58. Salazar O, Cirino PC, Arnold FH: **Thermostabilization of a cytochrome P450 peroxygenase**. *ChemBioChem* 2003, **4**:891-893.

59. Sakaue R, Kajiyama N: **Thermostabilization of bacterial fructosyl-amino acid oxidase by directed evolution**. *Appl Environ Microbiol* 2003, **69**:139-145.

60. Baik SH, Ide T, Yoshida H, Kagami O, Harayama S: **Significantly enhanced stability of glucose dehydrogenase by directed evolution**. *Appl Microbiol Biotechnol* 2003, **61**:329-335.

61. Oh KH, Nam SH, Kim HS: **Improvement of oxidative and thermostability of N-carbamyl-D-amino acid amidohydrolase by directed evolution**. *Protein Eng* 2002, **15**:689-695.

62. Murashima K, Kosugi A, Doi RH: **Thermostabilization of cellulosomal endoglucanase EngB from *Clostridium cellulovorans* by *in vitro* DNA recombination with non-cellulosomal endoglucanase EngD**. *Mol Microbiol* 2002, **45**:617-626.

63. Sriprapundh D, Vieille C, Zeikus JG: **Directed evolution of *Thermotoga neapolitana* xylose isomerase: high activity on glucose at low temperature and low pH**. *Protein Eng* 2003, **16**:683-690.

64. Waldo GS: **Genetic screens and directed evolution for protein solubility**. *Curr Opin Chem Biol* 2003, **7**:33-38.

65. Bulter T, Alcalde M, Sieber V, Meinhold P, Schlachtbauer C, Arnold FH: **Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution**. *Appl Environ Microbiol* 2003, **69**:987-995.